

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:50:19 ; Search time 30.0888 seconds  
(without alignments)  
2001.720 Million cell updates/sec

Title: US-10-004-219b-9

Perfect score: 2493

Sequence: 1 YNLICYFTNWAQYRGLGSF.....QQHQAGLVFTSCNCCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2089.5	83.8	455	22	AA62541 Disease treatment
2	2089.5	83.8	476	22	AA62543 Disease treatment
3	1662.5	66.7	368	23	ABB7580 Novel human protei
4	1387	55.6	377	22	AA62545 Disease treatment
5	1387	55.6	398	22	AA62544 Disease treatment
6	1301.5	52.2	466	18	AAW08584 Human 50 kDa chiti
7	1301.5	52.2	466	19	AAW40259 Human chitinase pr
8	1301.5	52.2	466	20	AAW42425 MO-218 clone of hu
9	1301.5	52.2	466	22	AAE00432 Human chitinase pr
10	1301.5	52.2	466	23	ABB76291 Human chitinase.

11	1295.5	52.0	466	19	AAW40260 Human chitinase pr
12	1295.5	52.0	466	20	AAW42426 MO-13B clone of hu
13	1295.5	52.0	466	22	AAE00433 Human chitinase pr
14	1295.5	52.0	466	23	ABB76292 Human chitinase.
15	1294.5	51.9	466	18	AAW31498 Human chitotrioid
16	1279.5	51.3	464	18	AAW31497 Human chitotrioid
17	1150.5	46.1	373	19	AAW40261 Human chitinase pr
18	1150.5	46.1	373	20	AAW42427 Clone of the C-ter
19	1150.5	46.1	373	23	ABB76293 Human chitinase tr
20	1148	46.0	373	19	AAW40262 Human chitinase pr
21	1148	46.0	373	20	AAW42428 Chitinase amino ac
22	1148	46.0	373	23	ABB76294 Human chitinase an
23	1146	46.0	387	18	AAW08585 Human 39 kDa chiti
24	1051	42.2	718	16	AAW33992 Murine oviduct spe
25	1042.5	41.8	668	16	AAW33993 Hamster oviduct sp
26	1010	40.5	537	16	AAW33991 Bovine oviduct spe
27	1008	40.4	696	22	ABBS8615 Drosophila melanog
28	962.5	38.6	383	21	ABBS8615 Gp38k protein sequ
29	956.5	38.4	383	18	AAW26751 Bovine whey protei
30	956.5	38.4	421	19	AAW53122 Amino acid sequenc
31	948.5	38.0	385	19	AAW53123 Amino acid sequenc
32	948.5	38.0	385	19	AAW47033 Human cartilage gp
33	948.5	38.0	416	19	AAW47034 Human cartilage gp
34	948.5	38.0	421	19	AAW53121 Amino acid sequenc
35	948.5	35.3	460	22	ABW35930 Human cartilage g1
36	879.5	35.3	460	22	ABW35930 Drosophila melanog
37	863.5	34.6	554	18	AAW01824 Manduca sexta larv
38	863.5	34.6	554	21	AAW07183 Manduca sexta gut
39	851	34.1	4498	22	ABBS8595 Drosophila melanog
40	782	31.4	595	22	ABW1737 Polypeptide involv
41	775	31.1	321	9	AAW81342 Human digestive sy
42	772.5	31.0	203	22	AAW92862 House dust mite (D
43	729.5	29.3	536	21	AAW52525 Der HMW-map poly
44	729.5	29.3	536	23	AAU96329 House dust mite (D
45	729.5	29.3	555	21	AAW52523 House dust mite (D

#### ALIGNMENTS

#### RESULT 1

AA62541  
ID AAG62541 standard; protein; 455 AA.  
XX  
AC AAG62541;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Disease treatment related protein SEQ ID NO: 1.  
XX  
DE Disease treatment; infection; chronic occlusive pulmonary disease;  
XX  
KW bronchial asthma.  
XX  
OS Homo sapiens.  
XX  
PN WO200136633-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 14-NOV-2000; 2000WO-JP08015.  
XX  
PR 15-NOV-1999; 99JP-0324467.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Nakanishi A, Morita S;  
XX  
DR WPI; 2001-397791/42.  
XX  
PT New proteins, peptides and DNA for treatment of bronchial asthma,  
XX  
PT chronic occlusive lung disease and infectious disease -  
XX  
PS Claim 1; Page 97-99; 114pp; Japanese.





PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;

PI WPI; 2002-292408/33.

XX DR N-PSDB; AEN32766.

XX An isolated polynucleotide for treating diseases associated with its

XX encoded polypeptide such as cancer and multiple sclerosis -

XX Claim 20; SEQ ID NO 848; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

XX novel human proteins. These were isolated from expressed sequences tags

XX (ESTs). They can be used to stimulate cell growth, to regulate

XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

XX e.g. in burn treatment, to regulate the immune system e.g. to treat

XX multiple sclerosis, to regulate activin or inhibit e.g. to treat

XX infertility, to regulate haemostasis or thrombolysis e.g. to treat

XX stroke and cancer, to screen for drugs, to treat inflammatory conditions

XX e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.

XX Parkinson's disease. The present sequence is a protein of the invention.

XX SQ Sequence 368 AA;

Query Match 66.7%; Score 1662.5; DB 23; Length 368;

Best Local Similarity 80.1%; Pred. No. 3.3e-137;

Matches 294; Conservative 35; Mismatches 35; Indels 3; Gaps 1;

QY 88 MYSTQNRQTFTSVIKFLRQYFGDGLDWEYPGSRGSPPODKHLFTVLVKEMREAFHQ 147

DB 1 MYSTPENRQTFTSVIKFLRQYFGDGLDWEYPGSRGSPPODKHLFTVLVKEMREAFHQ 60

QY 148 EAIENRRLMTAAVAGGISNQAGYEIPKSLYDLFHVMTYDLHGSWEYTGNSPL 207

DB 61 EAKQINKPLMTAAVAGGISNQAGYEIPKSLYDLFHVMTYDLHGSWEYTGNSPL 120

QY 208 YKYPTETGNSAYLVNDYVMYVWKNNGAPAEKLIIVGPPYGHFTFILRNPDSNGIGAPTSGD 267

DB 121 YKYPTDTGNSAYLVNDYVMYVWKNNGAPAEKLIIVGPPYGHFTFILRNPDSNGIGAPTSGA 180

QY 268 GPAGAYTQAGFWAYEICTFLRSAGTEVWDASQEVPPYAKANEMLYDNIKSFVKAQW 327

DB 181 GPAGPYAKESGIWAYEICTFLKNGATQGDAPQEVPPYQGVVWVGYDNIKSFIDIKAQW 240

QY 328 LKONNFGGAMIWAILDDFTGSCDQKPLTSTLNKALGISTEGCTAPDVPSEPVTTTP 387

DB 241 LKHNFPGGANWAILDDFTGFCQKPLSTLNKALGLOSACCTAPQPIETIAP 300

QY 388 ---GSGSGSGSGSGSGSGGFCADKADGLYPVADRNFAFWOCINGITYOQHCAQGLVFT 444

DB 301 SCSGNGSGSGSGSGSGSGGFCVAVRANGLYPVANNRNFAMHCVNGVITYQNCQAGLVFT 360

QY 445 SCNCCNW 451

DB 361 SCDCCNW 367

RESULT 4

AAG62545

ID AAG62545 standard; protein; 377 AA.

XX AC AAG62545;

XX DT 24-AUG-2001 (first entry)

XX DE Disease treatment related protein SEQ ID NO: 18.

XX

KW Disease treatment; infection; chronic occlusive pulmonary disease;

KW bronchial asthma.

OS Mus sp.

XX WO200136633-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-JP08015.

XX 15-NOV-1999; 99JP-0324467.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;

XX WPI; 2001-397791/42.

XX New proteins, peptides and DNA for treatment of bronchial asthma,

XX chronic occlusive lung disease and infectious disease -

XX Claim 18; Page 110-112; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be

XX used in the treatment and prevention of infectious diseases. Inhibitors

XX of the protein can be used to treat bronchial asthma and chronic

XX occlusive pulmonary disease. The present sequence is a protein described

XX in the exemplification of the invention.

XX SQ Sequence 377 AA;

Query Match 55.6%; Score 1387; DB 22; Length 377;

Best Local Similarity 66.2%; Pred. No. 5e-113;

Matches 249; Conservative 55; Mismatches 72; Indels 0; Gaps 0;

QY 1 YNLICYFTNWAQYRFLGSGFKPDDINPCLCTHLIYAFAGMQNNEITTIENNDVTLYKAFN 60

DB 1 YQLMCYFTSWAKDRPIEGSKFPGNIDPCLCTHLIYAFAGMQNNEITYTHEQDLRDYEALN 60

QY 61 DLKRNKSLKTLAIGGNWFTGAPTPTTWSTSONROTITSVIKFLRQYFGDGLDWEY 120

DB 61 GLKDKNTEKTLAIGGNWFTGAPTPTTWSTSONROTITSVIKFLRQYFGDGLDWEY 120

QY 121 PGRSGSPQDQKHLFTVLVKEMREAFQEAIESNRRLMTAAVAGGISNQAGYEIPKSL 180

DB 121 PGRSGSPQDQKHLFTVLVKEMREAFQEAIESNRRLMTAAVAGGISNQAGYEIPKSL 180

QY 181 KYLDFIHVMTYDLHGSWEYTGNSPLKYKPTETGNSAYLVNDYVMYVWKNNGAPAEKLI 240

DB 181 QSLDYIQWMTYDLHDPKDGTYGNSPLKSPYDICKSADLNVDISIYWKDHGAASEKLI 240

QY 241 VGFPYGHFTFILRNPDSNGIGAPTSGDGPAGAYTQAGFWAYEICTFLRSAGTEVWDAS 300

DB 241 VGFPYGHFTFILRNPDSNGIGAPTSGDGPAGAYTQAGFWAYEICTFLRSAGTEVWDAS 300

QY 301 QSPVPYAKANEMLYDNIKSFVKAQWLNKONNFGGAMIWAILDDFTGSCDQKPLTST 360

DB 301 QSPVPYAKANEMLYDNIKSFVKAQWLNKONNFGGAMIWAILDDFTGSCDQKPLTST 360

QY 361 TLNKALGISTEGCTAP 376

DB 361 TLNKALGISTEGCTAP 376

RESULT 5

AAG62544

ID AAG62544 standard; protein; 398 AA.

XX AC AAG62544;

XX DT 24-AUG-2001 (first entry)

XX

DE Disease treatment related protein SEQ ID NO: 17.  
XX Disease treatment; infection; chronic occlusive pulmonary disease;  
KW bronchial asthma.  
XX Mus sp.  
XX WO200136633-A1.  
XX 25-MAY-2001.  
XX 14-NOV-2000; 2000WO-JP08015.  
XX 15-NOV-1999; 99JP-0324467.  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX Nakanishi A, Morita S;  
XX WPI; 2001-397791/42.  
XX New proteins, peptides and DNA for treatment of bronchial asthma,  
PT -Chronic occlusive lung disease and infectious disease -  
XX Disclosure; Page 108-110; 114pp; Japanese.  
XX The present invention provides the sequence of a protein which can be  
CC used in the treatment and prevention of infectious diseases. Inhibitors  
CC of the protein can be used to treat bronchial asthma and chronic  
CC occlusive pulmonary disease. The present sequence is a protein described  
CC in the exemplification of the invention.  
XX Sequence 398 AA;  
SQ  
Query Match 55.6%; Score 1387; DB 22; Length 398;  
Best Local Similarity 66.2%; Pred. No. 5.4e-113;  
Matches 249; Conservative 55; Mismatches 72; Indels 0; Gaps 0;  
QY 1 YNLICVFTWAQVRPGLSFKPDDINPCLCTHLIYAFAGQNNNEITTIENNDVTLKAFN 60  
DB 22 YQLMCYTTWAKDRPIEGSFKNIDPCCLCTHLIYAFAGQNNNEITTIENNDVTLKAFN 81  
QY 61 DLKRNKSLKTLIAIGWNGFTAPFTTVMSTSONRQTFTITSVIFLRYQYCFDGLDLDWEY 120  
DB 82 GLKDKNTLKTLLAIGWNGFTAPFTTVMSTSONRQTFTITSVIFLRYQYCFDGLDLDWEY 141  
QY 121 PGSRGSPPQDHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNIQAGYEIPELS 180  
DB 142 PGSRGSPPQDHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNIQAGYEIPELS 201  
QY 181 KYLDFHVMYDHLHSGWEGYTGNSPLYKYPTETGNSNAYLVNDVYVMYKNGNAPAEKLI 240  
DB 202 QSLDYIQVMYDHLHSGWEGYTGNSPLYKYPTETGNSNAYLVNDVYVMYKNGNAPAEKLI 261  
QY 241 VGPPEYCHTFLRNPNDNGIGAPTSGDGPGAGVTRQAGFWAYYEICTFLRSGATEVWDAS 300  
DB 262 VGPPEYCHTFLRNPNDNGIGAPTSGDGPGAGVTRQAGFWAYYEICTFLRSGATEVWDAS 321  
QY 301 QEVPAKANEMLYGYNINISFVSVAQMLKONNFGGAMIWAILDDPTGFCQGGKPLPTS 360  
DB 322 QEVPAKANEMLYGYNINISFVSVAQMLKONNFGGAMIWAILDDPTGFCQGGKPLPTS 381  
QY 361 TLNKGALISTEGCTAP 376  
DB 382 TLNKGALISTEGCTAP 397

DT 24-MAR-1997 (first entry)  
XX Human 50 kDa chitinase.  
KW Chitinase; chitotriosidase; chitin; infectious disease;  
KW gene therapy; vaccine; diagnosis; lysosomal lipidosis;  
KW Gaucher disease; leishmaniasis; sarcoidosis; multiple sclerosis;  
KW X-linked adrenoleukodystrophy; drug delivery; cosmetics; food.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Key 1..21  
FT Peptide /label= Sig\_peptide  
FT Domain 178..198  
FT /label= Catalytic\_domain  
XX WO9640940-A2.  
XX 19-DEC-1996.  
XX 06-JUN-1996; 96WO-NL00225.  
XX 07-JUN-1995; 95US-0486839.  
XX (UNAM ) UNIV AMSTERDAM.  
XX Aerts JMG;  
XX WPI; 1997-118698/11.  
XX N-PSDB; AAT50833.  
XX New human chitinase - used to treat or prevent infection by  
PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,  
PT foods, implants etc.  
XX Claim 1; Page 40-42; 58pp; English.  
XX Human chitinases of 50 kDa (AAW08584) and 39 kDa (AAW08585) are stable  
CC to many proteases, active at pH 3-8 and up to 50 deg, and stable in  
CC the circulation. They are the result of alternative splicing of  
CC RNA, and can be produced on a large-scale in transformed host cells  
CC using cDNA clones (see also AAT50833-34) obtd. from a human macrophage  
CC library. The chitinases, optionally expressed from a gene therapy  
CC vector, are used to treat or prevent infection by chitin-contg.  
CC pathogens (e.g. fungi, protozoa, helminths) and more generally to  
CC degrade chitin. They are also useful in controlled-release drug  
CC delivery, in cosmetics, foods and dental products, for antibody  
CC prodn. and for diagnosis of diseases associated with elevated  
CC chitinase levels.  
XX Sequence 466 AA;  
SQ  
Query Match 52.2%; Score 1301.5; DB 18; Length 466;  
Best Local Similarity 51.5%; Pred. No. 2.1e-105;  
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;  
QY 3 LICYTNAQVRPGLSFKPDDINPCLCTHLIYAFAGQNNNEITTIENNDVTLKAFNDL 62  
DB 24 LVCYFTNAQVRPGLSFKPDDINPCLCTHLIYAFAGQNNNEITTIENNDVTLKAFNDL 83  
QY 63 KORNKSLKTLIAIGWNGFTAPFTTVMSTSONRQTFTITSVIFLRYQYCFDGLDLDWEYPG 122  
DB 84 KORNKSLKTLIAIGWNGFTAPFTTVMSTSONRQTFTITSVIFLRYQYCFDGLDLDWEYPG 143  
QY 123 SRGSPQDHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNIQAGYEIPELSKY 182  
DB 144 SQGSPAVDKERFTLLVQDLANAFQEAQTSQKERLLLSAAVPAGQTYVDAGYEVKIAQN 203  
QY 183 LDFHVMYDHLHSGWEGYTGNSPLYKYPTETGNSNAYLVNDVYVMYKNGNAPAEKLI 242  
DB 204 LDFVNLMAFYDHLHSGWEGYTGNSPLYKYPTETGNSNAYLVNDVYVMYKNGNAPAEKLI 263





Query Match	.52.2%	Score	1301.5;	DB	22;	Length	466;
Best Local Similarity	51.5%;	Pred. No.	2.1e-105;				
Matches	235;	Conservative	71;	Mismatches	129;	Indels	21;
Gaps	3						

  

QY	3	LICVFTNWAQYR	PCLGSKFKPDDINPCLCTHLI	YAPAGMNNETITIEWNDVTL	YKAFNDL	62	
DB	24	LVCIFTNWAQYR	QGEARFLP	KDLPDLPSLCTHLI	YAPAGTNHQLSTTEWNDVTL	YQEFNGL 83	
QY	63	KNRNSKLTKL	LAIAGGMNFGTAPFTT	MYSTSONROTFTITS	VIKFLROYGFDGLDLDWEYPG	122	
DB	84	KXNPKLTKL	LAIAGGMNFGTQKFT	DMVATANNROT	FVNSAIRFLRKYSF	DGLDLDWEYPG 143	
QY	123	SRGSPPODKHL	FTVLVKEMREAPQ	QEAIESNRPLMTAA	VAGGISNTOAGYEIPELSKY	182	
DB	144	SGSPSPAVDKER	FTTLVQDLANAFQ	QEAQTSGERLLLSAA	VPAGQTYVDAGYEV	DKIAQN 203	
QY	183	LDFTHWMTYDL	HSGWEGYTGENSE	PLYKYPTETGS	NAVILNDVVMNYWKN	NGAPAEKLVG 242	
DB	204	LDFNNLMAYD	PHGSEWKEVKTG	HNSEPLYKRBESG	AAASLNVDAAVQ	QMLQKGT	PASKLILG 263
QY	243	FPBYGHTFIL	RNPDSNDGIGAPT	SGDGPGAGYTRQAG	FWAYYEICTFL	RSGATEVMDASQE 302	
DB	264	MPYGRSFTL	ASSDTRVGAPAT	SGTGPPTKEG	MLAYEVCWS--	KGATKQRIQDK 321	
QY	303	VPYAYKAN	EWLYDNISF	SVKAWQLKONN	PGGAMIWAIDL	DDFTGSP	CDGKPKPLNSTL 362
DB	322	VPVIFRDNQ	WGFDDVESFKTK	VSYLKQKGLG	GMVWALDLD	DDFAGFS	CNQGRYPLIQTL 381
QY	363	NKALGISTE	CTAPDV-----	PSEPVTTP	PPSGSGSGSGSGSGSG	SGSGSGG	SGFCADKADGLYP 415
DB	382	QBELSLPYL	PSGTPELRV	PKPGQSEP-----	EHG	PSGQDTFCQ	KADGLYP 429
QY	416	VADDRNAFW	OCINGITYYQ	HCQAGLV	FDTSNC	CNCNW 451	
DB	430	NPRERSFY	SCAGRL	FQOCSCT	GLVFSN	CKCCTW 465	

  

RESULT 10	
ABB76291	
ID	ABB76291 standard; Protein; 466 AA.
XX	AC
XX	ABB76291;
XX	AC
DT	12-AUG-2002 (first entry)
XX	Human chitinase.
DE	
XX	
KW	Chitinase; enzyme; human; fungicide; antifungal; infection;
KW	candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
KW	paracoccidioidomycosis; histoplasmosis; cryptococcosis;
KW	chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
KW	Pneumocystis.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Location/Qualifiers
FT	1..21
FT	/label= Signal_peptide
FT	22..466
FT	/label= Mature_protein
XX	
PN	US6372212-B1.
XX	
PD	16-APR-2002.
XX	
XX	16-JUN-1997; 97US-0877599.
XX	
PR	14-JUN-1996; 96US-0663618.
XX	
PA	(ICOS-) ICOS CORP.
XX	
PI	Gray PW;

XX WPI; 2002-442449/47.  
DR N-PSDB; ABL57380.  
XX  
PT Co-administering chitinase to improve the effectiveness of fungicidal  
PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -  
XX  
XX Claim 1; Column 27-30; 26pp; English.  
XX  
XX The present sequence is the protein sequence of human chitinase  
CC as predicted from isolated cDNA clone MO-218 (see ABL57380). It  
CC differs in only 1 amino acid residue from the chitinase sequence  
CC (see ABB76292) deduced from a second cDNA clone, having glycine  
CC at position 81 of the mature protein. Northern blots showed  
CC highest chitinase gene expression in lung and ovary tissues.  
CC Expression in lung is consistent with a protective role against  
CC pathogenic organisms that contain chitin. The invention provides  
CC human chitinase polynucleotides and polypeptides, and materials and  
CC methods for the recombinant production of human chitinase products,  
CC which are expected to be useful as products for treating fungal  
CC infections or for the development of such products. Human  
CC chitinase has a synergistic effect on the actions of other  
CC fungicides. It can be administered to improve the antifungal  
CC activity of a non-chitinase antifungal agent, especially  
CC amphotericin B or itraconazole, in the treatment of a fungal  
CC infection such as candidiasis, aspergillosis, coccidioidomycosis,  
CC blastomycosis, paracoccidioidomycosis, histoplasmosis,  
CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,  
CC dermatophytoses and Pneumocystis infections (all claimed). In  
CC particular, the fungal infection involves candida, Aspergillus  
CC and/or Cryptococcus spp., whose growth is not effectively  
CC inhibited by contact with human chitinase alone.  
XX  
XX Sequence 466 AA;  
XX

Query Match 52.2%; Score 1301.5; DB 23; Length 466;  
Best Local Similarity 51.5%; Pred. No. 2.1e-105;  
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;  
3 LICFTNWAQYRGLGSKFPDDINPCLCTHLIIYAFAGMONEITTEWNVTLTKAFNDL 62  
24 LVCIYFNWAQYRGEARFLPKDLPDSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQENGL 83  
63 KNRNSKLKTLTLLAIGGNWFGTAPFTTWSTVSQNRQTFTTSVIKELROYGFDGLDLDWEYPG 122  
84 KKNWPKLTKTLTLLAIGGNWFGTKETDVAITANNRQTFVNSAIRFLRKYSPDGLDLDWEYPG 143  
123 SRGSPQDQKHLFTVLVKEMRAFEQBAIESNRRLMLVMTAAVAGISINQAGYEIPELSKY 182  
144 SQGSPAVDKERFTTLVQDLANAFQBEAQTSCKERLLLSAAVPAGQTVVDAGYEVDKIAQN 203  
183 LDFIHWMTVDLHGSWEGYTCENSPLYKYPTETGSNAYLVNDVYWNVKNNGAPAEKLVG 242  
204 LDFVNLMAVDFHGSWKBVTHGNSPLRYKQESGAASLNVDAAVQQWLQGTTPASKLILG 263  
243 FPEYGHFTILRNPDSNMGIGAPTSGDGPAGATYTRQAGFWAYYEICTFLRSGATEVWDASQE 302  
264 MPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGMLAYEYVCSW--KGATKQRIQDQK 321  
303 VPYAYKANEWLGYDNTKTSFVKAWMLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLSTL 362  
322 VPIYPRDNQWGVDFDYESPKTKVSYLKQKGLGAMVWALDDDFAGFSCNQGRYFLIQTLL 381  
363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGSGSGSGSGSGSGGCADKADGLYP 415  
382 RQELSLFYLPSGTPELEVPKPGQFSEF-----EHGFSPPGODTTCQKADGLYP 429  
416 VADDRNAFWQINGITVQHQCOAGLVFDTSCNCCNW 451  
430 NPRESRFSVCAAGRLFPQSCPTGLVFSNCKCCTW 465

```

RESULT 11
ID AAW40260
XX AAW40260 standard; Protein; 466 AA.
AC AAW40260;
XX
DT 15-JUN-1998 (first entry)
XX
DE Human chitinase protein from clone MO-13B.
XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= signal
FT /label= chitinase
FT /note= "from clone MO-13B"
XX
PN WO9747752-A1.
XX
PD 18-DEC-1997.
XX
PF 16-JUN-1997; 97WO-US10460.
XX
PR 14-JUN-1996; 96US-0663618.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW;
XX
XX WPI; 1998-052316/05.
DR N-PSDB; AAV10436.
XX
PT Nucleic acids encoding human chitinase - useful as antifungal
PT agents, especially in combination with other antifungals
XX
PS Claim 7; Page 44-45; 63pp; English.
XX
CC This sequence represents a novel human chitinase isolated from clone
CC MO-13B. Chitinases are useful for treating or preventing fungal infection
CC and as immunogens for generating antibodies which are used to purify,
CC detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.
CC The nucleic acid sequence of the chitinase is also useful as a probe to
CC identify and isolate genomic DNA encoding chitinases or similar proteins,
CC or cells expressing them or to generate transgenic ('knockout') rodents.
CC It can also be used in hybridisation assays and to detect genetic
CC alterations in the chitinase gene related to disease. Agents that inhibit
CC this protein may be useful in treatment of Gaucher's disease and
CC rheumatoid arthritis, where overexpression of the protein can damage
CC the extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,
CC and thus of their side effects.
XX
SQ Sequence 466 AA;

Query Match 52.0%; Score 1295.5; DB 19; Length 466;
Best Local Similarity 51.3%; Pred. No. 7.2e-105;
Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;

QY 3 LICYFTNWAQYRPGLSFPPDDINPCLTHLYAFAGMNNETIENNDVLYKAFNDL 62
DB 24 LVCYFTNWAQYRQGEARFLPKDLPCLTHLYAFAGMTNHLSTTWNDETLYQEFNGL 83
QY 63 KVRNSKLTLLAIGWNGFTAPFTTWTSTSONROTFTTSVIFLRCYGDGLDLWEYPG 122
DB 84 KQWPKLTLLAIGWNGFTQFTDVAFTANRQFTVNSAIRFLKYSFDGLDLWEYPG 143
QY 123 SRGSPQDKHFTLVVKEMREAFQEAIESNRPLMTVAATAVAGGISNIQAGYEIPELSKY 182

```

```

DB 144 SQGSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIAQN 203
QY 183 LDFIHWYTDLHGSWEGYTGENSPLYKYPTTETGNSAYLNVDYVMNYWKNNGAPAEKLIWG 242
DB 204 LDFVNLMAFYDFHGSWEKVTGHSNPLYKQEBESGAASLNVDAAVQOQWLQKGTTPASKLIIG 263
QY 243 PPEYGHTRILRNPSDNGIGAPTSGDPAGAVTROAGFWAYYEICTFLRSRGATEVWDASQE 302
DB 264 MPTYGRSTFLASSDTRVGATGSGTGPTKEGMLAYYEVCWS--KGATKQRIQDQK 321
QY 303 VPYAYKANEWLGYDNIKFSVKAQWLKQNNFGGAMIWALDLDFTGSCDQCKPFLTSTL 362
DB 322 VPIFRDNQWVGFDVSEFKTKVSYLKQKGLGGMVWALDLDPAFSGCNQGRYPLIOTL 381
QY 363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGSGSGSGSGSGGFCADKADGLYP 415
DB 382 RQELSLPYLPSTGTPLEVPKFPQPSPE-----EHGPFSGQDTFCQKADGLYP 429
QY 416 VADDRNATFQCCINGITYQOHOAGLVFDTSCNCCNW 451
DB 430 NPRESSPYSYCAAGRLFOQSCPTGLVFSNSCKCCTW 465

RESULT 12
AAV42426
ID AAY42426 standard; Protein; 466 AA.
XX
AC AAY42426;
XX
DT 10-DEC-1999 (first entry)
XX
DE MO-13B clone of human Chitinase; amino acid sequence.
XX
KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin-binding; allele; vector;
KW truncated protein; chitin binding region.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= Signal_peptide
FT /label= Mature_protein
XX
PN WO9946390-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US05343.
XX
PR 12-MAR-1998; 98US-0039198.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW, Tjoelker LW;
XX
DR WPI; 1999-551417/46.
DR N-PSDB; AAZ21848.
XX
PT Novel chitin-binding fragments of human chitinase used to treat fungal
PT infections in animals
XX
PS Example 1; Page 62-64; 83pp; English.
XX
CC This is the amino acid sequence of an allelic form of the human
CC chitinase enzyme, which is capable of degrading Chitin (a linear
CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).
CC The sequence also includes a Chitin binding peptide region, which lacks
CC any chitinase activity, but which can be used to target anti-fungal
CC agents to chitin containing fungal cells.
CC Chitinase fragments can be used to screen for proteins or other

```

CC molecules that specifically bind to the chitin-binding domain of human  
 CC chitinase or that modulate its activity. These compounds are useful for  
 CC immunization, as well as for purifying chitinase, as well as for  
 CC detection and quantification of chitinase. Polynucleotide fragments of  
 CC the invention are useful as a source of probes and primers, and to  
 CC express the proteins recombinantly. The chitinase fragments, when  
 CC conjugated to antifungal compounds, are used to treat animals,  
 CC especially humans, infected with chitin-containing parasites such as  
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
 CC sporotrichosis, and dermatophytoses.  
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
 CC protein for treating infections, especially fungal infections, is  
 CC problematic. In view of the increasing incidents of life-threatening  
 CC fungal infection in e.g. immunocompromised individuals, there exists a  
 CC need for identifying new compounds for treating fungal infection. The  
 CC chitin-binding fragments of the present invention provide this need.  
 XX  
 XX Sequence 466 AA;

Query Match 52.0%; Score 1295.5; DB 20; Length 466;  
 Best Local Similarity 51.3%; Pred. No. 7.2e-105;  
 Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;

QY 3 LICFTNWAQVRPGGLGSPKPDINPCLCTHLYAFAGMNNETIENWDTLYKAFNDL 62  
 DB 24 LVCYFTNWAQVRPGGLGSPKPDINPCLCTHLYAFAGMNNETIENWDTLYKAFNDL 83

QY 63 KNRNSKLTLLAIGGNWFTGAPFTTMTSTONRQFTITSVKFLRQYFGDGLDLDWEYPG 122  
 DB 84 KMNPKLTLLAIGGNWFTGAPFTTMTSTONRQFTITSVKFLRQYFGDGLDLDWEYPG 143

QY 123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNTQAGYEIPELSKY 182  
 DB 144 SQGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYEVNDKIAQN 203

QY 183 LDFIHWMTYDLHGSWEVGTGNSPLYKPYPTGTGSNAYLNVDYVMYWKNGAPASKLIVG 242  
 DB 204 LDFVNLMAFYDFHGSWEKVTGNSPLYKPYPTGTGSNAYLNVDYVMYWKNGAPASKLIVG 263

QY 243 PFYGHFTILRNPDSNGIGAPTSQDGPAGAYTRQAGFWAYEICTFLRSGATEVNDASQE 302  
 DB 264 MPTYGRSFTLASSDTRVGPATGSGTPGPTKEGGLAYEVCWS--KGATKQRIQDOK 321

QY 303 VPYAYKANWGLYDNIKFSVKAQWLKQNNFGAMWALDDFTGSCDQCKPFLITSTL 362  
 DB 322 VFYIFRDNQWGFDDVESFKTKVSLKQKGLGAMVWALDDFTGSCDQCKPFLITSTL 381

QY 363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGSGSGSGSGSGGFCADKADGLYP 415  
 DB 382 RQELSLPYLPSTPTELEVPKQPSPEP-----EHSQSPQDPTFCCKADGLYP 429

QY 416 VADDRNAPMQCINGITYQHQCOAGLVFTDSCNCNM 451  
 DB 430 NPERSSFYSCAAGRLFQSCPTGLVFSNCKCTW 465

RESULT 13

ID AAE00433

XX

AC AAE00433;

XX

DT 19-JUN-2001 (first entry)

XX

DE Human chitinase protein from clone pMO-13B.

XX

KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;

KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;

KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;

KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;

KW clone pMO-13B.

XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT Protein /label= Signal\_peptide  
 FT /label= Human\_mature\_chitinase\_protein  
 FT Domain 418..466  
 FT /label= Chitin\_binding\_domain  
 FT Region 1..373  
 FT /note= "Region with triacetylchitotriose  
 FT hydrolysing activity"  
 XX  
 XX WO200123430-A2.  
 XX  
 XX 05-APR-2001.  
 PD  
 XX 28-SEP-2000; 2000WO-US26960.  
 XX  
 XX 30-SEP-1999; 99US-0409918.  
 XX  
 XX (ICOS-) ICOS CORP.  
 PA  
 XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;  
 PI  
 XX WPI; 2001-266141/27.  
 DR  
 DR N-PSDB; AAD03760.  
 XX  
 PT Novel chitinase immunoglobulin fusion product, useful for treating  
 PT fungal infections and reducing the amount of a non-chitinase antifungal  
 PT agent needed for the treatment -  
 XX  
 PS Claim 1; Page 36-38; 39pp; English.  
 XX  
 CC The present invention relates to a chitinase immunoglobulin (Ig) fusion  
 CC product, comprising a human chitinase fused to at least a portion of an  
 CC immunoglobulin chain. The fusion product is useful for treating fungal  
 CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,  
 CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
 CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.  
 CC The fusion protein is useful for reducing the amount of non-chitinase  
 CC antifungal agent needed to exert an antifungal activity. The fusion  
 CC protein is also useful for preparing  
 CC a medicament for the prophylactic or therapeutic treatment of fungal  
 CC infections. Chitinase immunoglobulin fusion product has unexpectedly  
 CC improved serum half-life and formulation properties.  
 CC The present sequence is human chitinase protein from clone pMO-13B.  
 CC Chitinase enzyme degrades chitin which is a homopolymer of  
 CC beta-(1,4)-linked N-acetylglucosamine residues.  
 XX  
 SQ Sequence 466 AA;

Query Match 52.0%; Score 1295.5; DB 22; Length 466;  
 Best Local Similarity 51.3%; Pred. No. 7.2e-105;  
 Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;

QY 3 LICFTNWAQVRPGGLGSPKPDINPCLCTHLYAFAGMNNETIENWDTLYKAFNDL 62  
 DB 24 LVCYFTNWAQVRPGGLGSPKPDINPCLCTHLYAFAGMNNETIENWDTLYKAFNDL 83

QY 63 KNRNSKLTLLAIGGNWFTGAPFTTMTSTONRQFTITSVKFLRQYFGDGLDLDWEYPG 122  
 DB 84 KMNPKLTLLAIGGNWFTGAPFTTMTSTONRQFTITSVKFLRQYFGDGLDLDWEYPG 143

QY 123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNTQAGYEIPELSKY 182  
 DB 144 SQGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYEVNDKIAQN 203

QY 183 LDFIHWMTYDLHGSWEVGTGNSPLYKPYPTGTGSNAYLNVDYVMYWKNGAPASKLIVG 242  
 DB 204 LDFVNLMAFYDFHGSWEKVTGNSPLYKPYPTGTGSNAYLNVDYVMYWKNGAPASKLIVG 263





Search completed: June 29, 2003, 20:59:39  
Job time : 32.0888 secs

09-OCT-1997.  
21-MAR-1997; 97WO-US05072.  
29-MAR-1996; 96US-0014295.  
(HUMA-) HUMAN GENOME SCI INC.  
(SMIK ) SMITHKLINE BEECHAM CORP.  
Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR,  
WPI; 1997-503041/46.  
N-PSDB; AAT89181.  
New isolated human chitotriosidase gene - used to develop products  
for the diagnosis and treatment of tissue remodeling disorders, e.g.  
rheumatoid arthritis  
Disclosure; Page 22-24; 34pp; English.  
This protein comprises a variant of human chitotriosidase (see  
also AA031497). Its amino acid sequence was deduced from the  
coding sequence of a full-length cDNA clone (see AAT89181). The  
following are claimed: (1) a nucleic acid sequence encoding  
chitotriosidase; (2) a method of diagnosing a tissue remodeling  
disorder related to expression of a mutated chitotriosidase protein  
in a host comprising carrying out nucleic acid amplification; and  
CC. (3) a method of detecting altered expression of a chitotriosidase  
protein in a host comprising contacting a bodily sample with an  
antibody. Chitotriosidase can degrade extracellular matrix  
substrates with a similar carbohydrate structure to chitin. It can  
be used to develop products which can be used in the diagnosis and  
treatment of tissue remodeling disorders such as rheumatoid  
arthritis or atherosclerosis.

Query Match 51.9%; Score 1294.5; DB 18; Length 466;  
Best Local Similarity 51.3%; Pred. No. 8.8e-105;  
Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;  
SQ Sequence 466 AA;  
3 LICYFTNNAQYRPGKPGDDINCLCTHLIYAFAGMONEITWNVDTLYKAPNDL 62  
24 LVCYFTNNAQYRQGEARFLPKDLPDSLCTHLIYAFAGMTHQSLTWNDETLYQEFNGL 83  
63 KNENSKLITLLAIGCWNFTAPFTTWSTSONROTFTSVIKFQYGFGLDLDWEYPG 122  
84 KQNPPLKLTLLAIGGWNFTQKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEYPG 143  
123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGGISNIOAGYEIPELSKY 182  
144 SGSPAVDKERTTLLVQDLANAFQEAQTSKXERLLLSAIVPAGQTYVDAGYEVDKIAQN 203  
183 LDFIHMVYDLHGSNEGVTGENSPLYKYPTETGSNAYLNVDYVMYKNGAPAEKLIWG 242  
204 LDFVNLMAVDPHGSEKVTGHNSPLYKQESGAAASLNVDAAVQWLQKGTSPASKLILG 263  
243 PPEYGHFTILRNPSDNGIGAPTSGDGAGAYTRQAGFWAYYEICTFLRSGATEVWDASQE 302  
264 MFTYGRSRFTLASSSTRVGPATGSGTGPFTKEGGLAYYEVCWSW--KGATKQRIQDK 321  
303 VPYAYKANEWLYGDIKGSFVKAQWLKONNFGGAMIWALDLDFTGSCDQCKFPLTSTL 362  
322 VPIYFRDQWGVDFVESFKTVSVLKQKGLGGAMVWALDLDLDFAGFCNQGRIYLIQTL 381  
363 NKALGISTEGCTAPDV-----PSEPVTPPGSGGGSGSGSGGFCADKADGLYP 415  
382 RQELSLPVLPSGTPELVKPGQPSSEP-----EHGFPSPQDTFCQKADGLYP 429  
416 VADDRNATFWQINGITYQHQOAGLVFDTSCNCCNW 451  
430 NPRERSFSYCAAGRLFQGCPTGLVFNNSCKCCTW 465

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:57:59 ; Search time 11.7012 seconds  
(without alignments)  
1136.565 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLICYFTNWAQYRPLGSLF.....QQHCQAGLVFTSCNCCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1301.5	52.2	466	2	US-08-486-839-4	Sequence 4, Appli
2	1301.5	52.2	466	3	US-09-151-011-4	Sequence 4, Appli
3	1301.5	52.2	466	4	US-09-039-198A-2	Sequence 2, Appli
4	1301.5	52.2	466	4	US-09-343-623-4	Sequence 4, Appli
5	1301.5	52.2	466	4	US-08-877-599-2	Sequence 2, Appli
6	1301.5	52.2	466	4	US-09-267-574-2	Sequence 2, Appli
7	1295.5	52.0	466	4	US-09-039-198A-4	Sequence 4, Appli
8	1295.5	52.0	466	4	US-08-877-599-4	Sequence 4, Appli
9	1295.5	52.0	466	4	US-09-267-574-4	Sequence 4, Appli
10	1150.5	46.1	373	4	US-09-039-198A-14	Sequence 14, Appl
11	1150.5	46.1	373	4	US-08-877-599-14	Sequence 14, Appl
12	1150.5	46.1	373	4	US-09-267-574-14	Sequence 14, Appl
13	1148	46.0	373	4	US-09-039-198A-15	Sequence 15, Appl
14	1148	46.0	373	4	US-08-877-599-15	Sequence 15, Appl
15	1148	46.0	373	4	US-09-267-574-15	Sequence 15, Appl
16	1146	46.0	387	3	US-08-486-839-6	Sequence 6, Appli
17	1146	46.0	387	3	US-09-151-011-6	Sequence 6, Appli
18	1146	46.0	387	4	US-09-343-623-6	Sequence 6, Appli
19	973.5	39.0	383	4	US-09-459-749D-17	Sequence 17, Appl
20	948.5	38.0	385	2	US-08-694-915-2	Sequence 2, Appli
21	948.5	38.0	416	2	US-08-694-915-4	Sequence 4, Appli
22	863.5	34.6	554	2	US-08-524-051-2	Sequence 2, Appli
23	863.5	34.6	554	3	US-09-052-778-16	Sequence 16, Appl
24	802	32.2	559	4	US-09-545-814-14	Sequence 14, Appl
25	802	32.2	583	4	US-09-545-814-2	Sequence 2, Appli
26	802	32.2	583	4	US-09-545-814-5	Sequence 5, Appli
27	799	32.0	635	4	US-09-545-814-32	Sequence 32, Appl

28	729.5	29.3	536	4	US-09-292-225-21	Sequence 21, Appl
29	729.5	29.3	555	4	US-09-292-225-15	Sequence 15, Appl
30	729.5	29.3	555	4	US-09-292-225-18	Sequence 18, Appl
31	722	29.0	490	4	US-09-292-225-41	Sequence 41, Appl
32	722	29.0	509	4	US-09-292-225-35	Sequence 35, Appl
33	722	29.0	509	4	US-09-292-225-38	Sequence 38, Appl
34	667.5	26.8	489	4	US-09-545-814-29	Sequence 29, Appl
35	470.5	18.9	389	1	US-07-939-501A-1	Sequence 1, Appli
36	470.5	18.9	389	4	US-08-448-398-7	Sequence 7, Appli
37	470.5	18.9	423	1	US-07-939-501A-10	Sequence 10, Appl
38	470.5	18.9	423	1	US-07-939-501A-12	Sequence 12, Appl
39	455.5	18.3	424	1	US-08-045-289C-2	Sequence 2, Appli
40	455.5	18.3	424	3	US-08-371-680-2	Sequence 2, Appli
41	455.5	18.3	424	3	PCT-US94-01198-2	Sequence 2, Appli
42	434	17.4	442	3	US-09-052-778-2	Sequence 2, Appli
43	418	16.8	371	2	US-08-591-629-2	Sequence 2, Appli
44	409	16.4	377	2	US-08-591-629-8	Sequence 8, Appli
45	389	15.6	452	3	US-09-052-778-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-08-486-839-4  
; Sequence 4, Application US/08486839  
; Patent No. 5928928  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A human chitinase, its recombinant  
; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron  
; STREET: 350 Gericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11758  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,839  
; FILING DATE: 07 - June - 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baron, Ronald J.  
; REGISTRATION NUMBER: 29,281  
; REFERENCE/DOCKET NUMBER: 294-26  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-08-486-839-4

Query Match 52.2%; Score 1301.5; DB 2; Length 466;  
Best Local Similarity 51.5%; Pred. No. 1.4e-108;  
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;  
QY 3 LICYTNWAQYRPLGSLFKPDDINPCLCTHLYAFAGHONNEITTIWVDVLYKAFNDL 62  
DB 24 LVCYTNWAQYRQSGARFLPKDLPCLCTHLYAFAGHONNEITTIWVDVLYKAFNDL 83

[illegible]

RESULT 2  
 US-09-151-011-4  
 ; Sequence 4, Application US/09151011  
 ; Patent No. 6057142  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: A Human Chitinase, Its Recombinant  
 ; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in  
 ; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hoffmann & Baron, LLP  
 ; STREET: 6900 Jericho Turnpike  
 ; CITY: Syosset  
 ; STATE: New York  
 ; COUNTRY: United States of America  
 ; ZIP: 11791  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; SOFTWARE:  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/151.011  
 ; FILING DATE: 10 - September - 1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Morris, Robert C.  
 ; REGISTRATION NUMBER: 42,910  
 ; REFERENCE/DOCKET NUMBER: 294-32 DIV  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (516) 822-3550  
 ; TELEFAX: (516) 822-3582  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 466 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; US-09-151-011-4

```

Query Match          52.2%; Score 1301.5; DB 3; Length 466;
Best Local Similarity 51.5%; Pred. No. 1.4e-108;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3

QY      3  LICVFTNWAQYRPGLSFKFDDINPCLCTHLIIYAFAGMNNETITWNDVTLTKAFNDL 62
DB      24  LVCVFTNWAQYRQGEARFLPKDLDPLCTHLIIYAFAGMTNHLSTWMDLTLYQEFNGL 83

QY      63  KNRSKLTLLAIGWNFGTAPTITWVSTSONRQTEITSVIKFLROYGPDGLDLWDWEYPG 122
DB      84  KWMFKLTLLAIGWNFGTQKTDWVATANNRQTFVNSAIRFLKYSFDGLDLWDWEYPG 143

QY      123  SRGSPPOKHLFTVLVKEMREAFEQEBAIESNRPLMVTAAVAGGISNIOAQYVEIPELSKY 182
DB      144  SQSPAVDWKERFTTLVQDLANAFQGEAQTSKBERLLLSAAVPAGQTYVDAGYEVVDKIAQN 203

QY      193  LDFTHVMTYDLHGSWEQYTGENSEPLKYKPYETCSNAYLNVYVMYKKNNGAPAKLIUVG 242
DB      204  LDFVNLWAYDPHGSWEKVTGHNFLYKQEEGAAASLNVDAVQOMLQKGTTPASKLILG 263

QY      243  PPEYCHTFILRNPDNGIGAPTSGDGPAGATYRQAGFWAYYIEICTFLRSGATEVWDASQE 302
DB      264  MPTVGRSFTLASSSDTRVGAPATGSGTPGPFKEGGMILAYYECVW--KGATKQRIQDQK 321

QY      303  VPYAYKANEMLYGDNIKSFVSKAQWLKONNFGGAMITWAILDDFTGSPCDQGKFLTSTL 362
DB      322  VPIYFRDNQWVGFPDVBESFKTKVSYLKQGLGGAMVWALDLDLDDFAGFSCNQGRYFLIQT 381

QY      363  NKALIGISTEGCTADPV-----PSPVPTTTPGSGSGSGSGSGSGSGSGGFCADKADGLYP 415
DB      382  RQELSLPLPSGTPELEVPKPGQFSEP-----EHGSPSPQDFTFCQKADGLYP 429

QY      416  VADNRNFMWOCINGITVQOHCQAGLVFDTSCNCNW 451
DB      430  NPRESSFYSCAAGRLPQQSCPTGLVFSNCKCCTW 465

RESULT 3
US-09-039-198A-2
; Sequence 2, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid

```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-039-198A-2

Query Match      52.2%; Score 1301.5; DB 4; Length 466;
Best Local Similarity 51.5%; Pred. No. 1.4e-108;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LICYFTNWAQYRPGKSPKDDINPCLCTHLYAFAGMNNIITIEWNDVLYKAFNDL 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 24 LVCYFTNWAQYRQGEARFLPKDLPCLCTHLYAFAGMNNHOLSTTEWDETLYQEFNGL 83
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 63 KRNKSLKTLAIGGNWFGTAPFTTWSTQNRQFTITSVIFKRYQFGDGLDWEYPG 122
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 84 KKNPKLKTLLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLKYSFGDGLDWEYPG 143
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 123 SRGSPDQKHLFTVLVKEMREAFQEAIESNRPRLMTVAAGGISNQAGYEIPELSKY 182
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 144 SQGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQOTYVDAGYEVDKIAQN 203
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 183 LDFIHMTYDLHGSWEGYTGNSPLYKYPTETGNSNAYLNVDYVMYKNGGAPAEKLIYG 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 204 LDFVNLMAIDFHGSWEKVTGHSNPLYKQESGAASLNVDAAVQWLQKGTPTASKLIIG 263
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 243 FPEYGHFTILRNPSDNGIGAPTSGDGAGAYTRQAGFWAYYIEICTFLRSGATEVWDASOE 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 264 MPTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGLAYYEVCSW--KGATKQRIQDQK 321
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 303 VPYAYKANWELGYDNIKSPVKAQWLKONNFGGAMIWALDDFTGSPCDQCKFPLTSTL 362
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 322 VPYIFRDNQWGFDDVESFKTKVSLKOKGLGGMVWALDDDFAGFCNCQGRYPLIQTLL 381
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGGGSGSGSGGSCFADKADGLYP 415
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 382 RQELSLPLPSGTPELEVPKQCPSEP-----EHGSPSQDFTFCQKADGLYP 429
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 416 VADDRNAPWQINGITYQOHCQAGLVFDTSCNCCNW 451
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 430 NPRESSFSYSCAAGRLFQSCPTGLVFSNSCKCCTW 465
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 4
US-09-343-623-4
; Sequence 4, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281

;
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-343-623-4

Query Match      52.2%; Score 1301.5; DB 4; Length 466;
Best Local Similarity 51.5%; Pred. No. 1.4e-108;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LICYFTNWAQYRPGKSPKDDINPCLCTHLYAFAGMNNIITIEWNDVLYKAFNDL 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 24 LVCYFTNWAQYRQGEARFLPKDLPCLCTHLYAFAGMNNHOLSTTEWDETLYQEFNGL 83
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 63 KRNKSLKTLAIGGNWFGTAPFTTWSTQNRQFTITSVIFKRYQFGDGLDWEYPG 122
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 84 KKNPKLKTLLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLKYSFGDGLDWEYPG 143
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 123 SRGSPDQKHLFTVLVKEMREAFQEAIESNRPRLMTVAAGGISNQAGYEIPELSKY 182
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 144 SQGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQOTYVDAGYEVDKIAQN 203
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 183 LDFIHMTYDLHGSWEGYTGNSPLYKYPTETGNSNAYLNVDYVMYKNGGAPAEKLIYG 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 204 LDFVNLMAIDFHGSWEKVTGHSNPLYKQESGAASLNVDAAVQWLQKGTPTASKLIIG 263
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 243 FPEYGHFTILRNPSDNGIGAPTSGDGAGAYTRQAGFWAYYIEICTFLRSGATEVWDASOE 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 264 MPTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGLAYYEVCSW--KGATKQRIQDQK 321
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 303 VPYAYKANWELGYDNIKSPVKAQWLKONNFGGAMIWALDDFTGSPCDQCKFPLTSTL 362
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 322 VPYIFRDNQWGFDDVESFKTKVSLKOKGLGGMVWALDDDFAGFCNCQGRYPLIQTLL 381
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGGGSGSGSGGSCFADKADGLYP 415
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 382 RQELSLPLPSGTPELEVPKQCPSEP-----EHGSPSQDFTFCQKADGLYP 429
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 416 VADDRNAPWQINGITYQOHCQAGLVFDTSCNCCNW 451
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 430 NPRESSFSYSCAAGRLFQSCPTGLVFSNSCKCCTW 465
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 5
US-08-877-599-2
; Sequence 2, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
```

```

; SEQ ID NO 2
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-267-574-2

Query Match          52.2%; Score 1301.5; DB 4; Length 466;
Best Local Similarity 51.5%; Pred. No. 1.4e-108;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LICFTNWAQYRPGLSFKPDDINPCLC HL |YAFAGMNNETITIEWNDVLYKAFNDL 62
DB 14 LVCYPTNWAQYRGEARPLPKDLSCL TH |LYAFAGMTNHLSTTEWMDVLYQSFNGL 83
QY 63 KURNKSLKTLAIGWNFGTAFTHWVSTSONROTFTITSVIKFLQYGFGLDLDWEYFG 122
DB 84 KQNNPKLTLAIGWNFGTQFTDMVATANNRQTFVNSAIRFLRKYSFDGLDLDWEYFG 143
QY 123 SRGSPQDQKHLFTVLVKENREAFQEAIESNPRLMVTAAVAGGISNIOAGYEIPELSKY 182
DB 144 SOGSPAVDKERTTLVQDLANAFQEAQTSKGERLLLSAAVAGQTYVDAGVEVDKIAQN 203
QY 183 LDFIHMTYDLHSGMEGYTGNSPLYKYKPYETGTGSNAYLNVYVMYVKNKGAPAEKLIWG 242
DB 204 LDFVNLWAYDFHSGWEKVGTGHNSPLYKROEESGAAASLVNDAVQOVLQKGTSPASKLIIG 263
QY 243 FPEYGHFTILRNPSONGIGAPTSGDGPAGAYTRQAGFWAYEYICITFLRSGSEVMDASQ 302
DB 264 MPTYGRSFTLASSDTRVGPATGSGTPGFTFKGGMLAYEVCW--KGATKQRIQDQK 321
QY 303 VPYAYKANEWLGNDYIKSFVSKAQWLKQNNFGGAMIWALDLDFTGSCFCQCKFPLTSTL 362
DB 322 VFYIFRDNQWGFDDVESPKTKVSYLKQKGLGGAMVWALDLDFTGSCFCQCKFPLTSTL 381
QY 363 NKALGISTEGCTAPV-----PSEPVTTPGSSGGSSGGSSGGSSGGSCFCADKADGLYP 415
DB 382 ROELSLPYLPDGTPELVEPKQPSEP-----EHPGSPGQDTFCQCKADGLYP 429
QY 416 VADDRNFAWQCINGITYQQHCOAGLVFDTSCNCCNW 451
DB 430 NPERSSFFVSCAAGRLFQSCPTGLVFSNCKCCTW 465

RESULT 7
US-09-039-198A-4
; Sequence 4, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marehall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:

```

```

RESULT 9
US-09-267-574--4
; Sequence 4, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 4
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-09-267-574-4

Query Match 52.0%; Score 1295.5; DB 4; Length 466;  
Best Local Similarity 51.3%; Pred. No. 4.7e-108;  
Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;

QY 3 LICYFTNWAQYRPGSGFKPDDINPCLCTHLIYAFAGMONEITTIEMNDVTLKAFNDL 62  
DB 24 LVCYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGMNHQSLSTTEWDETLYQEFNGL 83  
QY 63 KNRNSKLTLLAIGWNFGTAPFTTMVSTSONROTFTITSVIFKFLROYGDFGLDLDWEYPG 122  
DB 84 KQNPPLKLTLLAIGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEYPG 143  
QY 123 SRGSPQDQKHLFTVLVKEMREAFQEAIESNRPLMVTAAVAGGISNIQAGYEIPELSKY 182  
DB 144 SQSPAVDKERTFTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYEVDKIAQ 203  
QY 183 LDFHVMYDLHSGWEGYTGNSPLYKYPTETGSAVYLVNVDYVMYWKNGAPAEKLI 242  
DB 204 LDFVNLMAVDFHSGWEKVTGHSNPLYKQEESSAAASLNVDAVQVQLQKGTTPASKLI 263  
QY 243 PFEYGHFTILNPSDMGICGAPTSQDPAGATYQAGFWAYYEICTFLRSGATEVWDASOE 302  
DB 264 MPTYGRSFTLASSSDTRVGPATGSGTGPFTKEGGMLAYYEVCSW--KGATKQRIQDOK 321  
QY 303 VPYAYKANEMWGLVDNIKSFVKAQWLKONNFGGAMIWADLDFTGSCDQKFPPLTSTL 362  
DB 322 VPIYFDNQWGFDDVESFKTKVSLKQKGLGGMVWALDLDFFAGFCNQGRYPLIQT 381  
QY 363 NKALGISTEGCTAPDV-----PSEPTVTPPGSGSGSGSGSGSGSGGFCADKADGLYP 415  
DB 382 RQELSLPYLPGTPELEVPKPGQSEP-----EHGSPQDFTFCQKADGLYP 429  
QY 416 VADDRNFWQCINGITYQHOCQAGLVFDTSCNCCNW 451  
DB 430 NPRRSSFYSCAAGRLFOQSCPTGLVFNSSRCCTW 465

RESULT 10

US-09-039-198A-14  
Sequence 14, Application US/09039198A  
Patent No. 6200951  
GENERAL INFORMATION:  
APPLICANT: Gray, Patrick W.  
APPLICANT: Tjoelker, Larry W.  
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,198A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 27866/34391  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-039-198A-14

Query Match 46.1%; Score 1150.5; DB 4; Length 373;  
Best Local Similarity 54.2%; Pred. No. 3.6e-95;  
Matches 205; Conservative 65; Mismatches 99; Indels 9; Gaps 2;

QY 3 LICYFTNWAQYRPGSGFKPDDINPCLCTHLIYAFAGMONEITTIEMNDVTLKAFNDL 62  
DB 3 LVCYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGMNHQSLSTTEWDETLYQEFNGL 62  
QY 63 KNRNSKLTLLAIGWNFGTAPFTTMVSTSONROTFTITSVIFKFLROYGDFGLDLDWEYPG 122  
DB 63 KQNPPLKLTLLAIGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEYPG 122  
QY 123 SRGSPQDQKHLFTVLVKEMREAFQEAIESNRPLMVTAAVAGGISNIQAGYEIPELSKY 182  
DB 123 SQSPAVDKERTFTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYEVDKIAQ 182  
QY 183 LDFHVMYDLHSGWEGYTGNSPLYKYPTETGSAVYLVNVDYVMYWKNGAPAEKLI 242  
DB 183 LDFVNLMAVDFHSGWEKVTGHSNPLYKQEESSAAASLNVDAVQVQLQKGTTPASKLI 242  
QY 243 PFEYGHFTILNPSDMGICGAPTSQDPAGATYQAGFWAYYEICTFLRSGATEVWDASOE 302  
DB 243 MPTYGRSFTLASSSDTRVGPATGSGTGPFTKEGGMLAYYEVCSW--KGATKQRIQDOK 300  
QY 303 VPYAYKANEMWGLVDNIKSFVKAQWLKONNFGGAMIWADLDFTGSCDQKFPPLTSTL 362  
DB 301 VPIYFDNQWGFDDVESFKTKVSLKQKGLGGMVWALDLDFFAGFCNQGRYPLIQT 360  
QY 363 NKALGISTEGCTAPDVPS 380  
DB 361 RQELSL-----PYLPS 371

RESULT 11

US-08-877-599-14  
Sequence 14, Application US/08877599  
Patent No. 6372212  
GENERAL INFORMATION:  
APPLICANT: Gray, Patrick W.  
TITLE OF INVENTION: Chitinase Materials and Methods  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/877,599  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/663,618  
FILING DATE: 14-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 27866/33994  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300



TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-877-599-14

Query Match 46.1%; Score 1150.5; DB 4; Length 373;  
Best Local Similarity 54.2%; Pred. No. 3.6e-95;  
Matches 205; Conservative 65; Mismatches 99; Indels 9; Gaps 2;

QY 3 LICVYTNNAQVRPGSGFKPDDINPCLCTHLIYAFAGMONEITIEWNDVLYKAFNDL 62  
DB 3 LVCYFTNNAQVRQGEARFLPKDLPSCSLCTHLIYAFAGMTHOLSTEWNDVLYQEFNGL 62  
QY 63 KNRNSKLTLLAIGWNGFTAPFTTMVSTSQNRQTFITSVIKFLRQYFGDGLDLWEYPG 122  
DB 63 KMNPKLKTLLAIGWNGFTQKFTDMVATANNRQTFVNSAIRFLKYSFDFGLDLWEYPG 122  
QY 123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNIOAGVEIPELSKY 182  
DB 123 SQGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQOTYVDAGYEVVDKIAQN 182  
QY 183 LDFIHVMYDHLGSGWEGYTGNSPLYKYPTETGSGNAYLNVDVVMYKNGNGAPAEKLI 242  
DB 183 LDFVNLMAFDHSGWKEVYTGNSPLYKQESGAASLVNDAVQOVLQKGTTPASKLI 242  
QY 243 PPEYGHTEFLRNPDSNGIGAPTSGDPAGAYTROAGFWAYVEICTFLRSGATEVWDASQE 302  
DB 243 MPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGLAYVEVCWS--KGATKQRIQDOK 300  
QY 303 VPIYAYKANWLYGDNKISFVKAQWLKQNNFGGAMIWALDDFTGSCDQKFPPLTSTL 362  
DB 301 VPIYFRDNQWYGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAGFCNQGRIYPLIOTL 360  
QY 363 NKALGISTEGCTAPDVPS 380  
DB 361 RQELSL-----PYLPS 371

RESULT 12  
US-09-267-574-14  
Sequence 14, Application US/09267574  
Patent No. 6399571  
GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.  
APPLICANT: Tjoelker, Larry W.  
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
FILE REFERENCE: 27866/35407  
CURRENT APPLICATION NUMBER: US/09/267,574  
CURRENT FILING DATE: 1999-03-12  
EARLIER APPLICATION NUMBER: 09/039,198  
EARLIER FILING DATE: 1998-03-12  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 14  
LENGTH: 373  
TYPE: PRT  
ORGANISM: Homo sapiens

US-09-267-574-14

Query Match 46.1%; Score 1150.5; DB 4; Length 373;  
Best Local Similarity 54.2%; Pred. No. 3.6e-95;  
Matches 205; Conservative 65; Mismatches 99; Indels 9; Gaps 2;

QY 3 LICVYTNNAQVRPGSGFKPDDINPCLCTHLIYAFAGMONEITIEWNDVLYKAFNDL 62  
DB 3 LVCYFTNNAQVRQGEARFLPKDLPSCSLCTHLIYAFAGMTHOLSTEWNDVLYQEFNGL 62

QY 63 KNRNSKLTLLAIGWNGFTAPFTTMVSTSQNRQTFITSVIKFLRQYFGDGLDLWEYPG 122  
DB 63 KMNPKLKTLLAIGWNGFTQKFTDMVATANNRQTFVNSAIRFLKYSFDFGLDLWEYPG 122  
QY 123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNIOAGVEIPELSKY 182  
DB 123 SQGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQOTYVDAGYEVVDKIAQN 182  
QY 183 LDFIHVMYDHLGSGWEGYTGNSPLYKYPTETGSGNAYLNVDVVMYKNGNGAPAEKLI 242  
DB 183 LDFVNLMAFDHSGWKEVYTGNSPLYKQESGAASLVNDAVQOVLQKGTTPASKLI 242  
QY 243 PPEYGHTEFLRNPDSNGIGAPTSGDPAGAYTROAGFWAYVEICTFLRSGATEVWDASQE 302  
DB 243 MPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGLAYVEVCWS--KGATKQRIQDOK 300  
QY 303 VPIYAYKANWLYGDNKISFVKAQWLKQNNFGGAMIWALDDFTGSCDQKFPPLTSTL 362  
DB 301 VPIYFRDNQWYGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAGFCNQGRIYPLIOTL 360  
QY 363 NKALGISTEGCTAPDVPS 380  
DB 361 RQELSL-----PYLPS 371

#### RESULT 13

US-09-039-198A-15  
Sequence 15, Application US/09039198A  
Patent No. 6200951

GENERAL INFORMATION:  
APPLICANT: Gray, Patrick W.  
APPLICANT: Tjoelker, Larry W.  
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,198A  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 27866/34391  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-039-198A-15

Query Match 46.0%; Score 1148; DB 4; Length 373;  
Best Local Similarity 55.2%; Pred. No. 6e-95;  
Matches 202; Conservative 64; Mismatches 98; Indels 2; Gaps 1;

QY 3 LICVYTNNAQVRPGSGFKPDDINPCLCTHLIYAFAGMONEITIEWNDVLYKAFNDL 62  
DB 3 LVCYFTNNAQVRQGEARFLPKDLPSCSLCTHLIYAFAGMTHOLSTEWNDVLYQEFNGL 62

Qy	63	K0RNSKLTLLAIGGNFGTAPFTTWSTVSTQNRQTFTTSWIKELQYQFGDGLDWEYGP	122
Db	63	K0MNPKLTLAIGGNFGTKFTDMVATANNRQTFVNSAIRFLKYSFGDGLDWEYGP	122
Qy	123	SRGSPPODKHFTVLVKEMREAFQOEAIENRPLMWTAAVAGGISNIQAGYEIPELSKY	182
Db	123	SGSPSPAVDKERFTTLLVQDLANAFQOEAGTSGKERLLLSAAVPAGOTVVDAGYEVDDIAQN	182
Qy	183	LDFIHWMTYDLHSGMEGTGNSPLYKYPTETGNSAYLVNDVYMWNYWKNNGAPAEKLI	242
Db	183	LDFVNLMAVDFHSGMEKVTGNSPLYKKESGGAASLNVDAAVQOMLQKGTSPAKLI	242
Qy	243	FPEYGHFTFLRNPSONGICAPTSDDGPAGAYTRQAGFWAYYEICTFLRSGATEVMDASOE	302
Db	243	MPTYGRSFTLASSDTRVGAATGSGTPGPTKESGMLAYYEVCSW--KGATKQRLODQK	300
Qy	303	VPIYAYKANEWGLGYDNIKSFSVKAOMLKQNNPGGAMIWAIDLDDPTGSCFDQKGFPTSTL	362
Db	301	VPVYPRDNQWGVDFDVESEFKTKVSYLKQKGLGGAMVWALDLDDFAGFSCNQGRPYLIQTL	360
Qy	363	NKALGI 368	
Db	361	ROELSL 366	

```

RESULT 14
US-08-877-599-15
; Sequence 15, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-877-599-15

Query Match 46.0%; Score 1148; DB 4; Length 373;
Best Local Similarity 55.2%; Pred. No. 6e-95;
Matches 202; Conservative 64; Mismatches 98; Indels 12; Gaps 1

```

Qy	3	LICYTNNWAQYRPGLGSPKDDINPCLCTHLIIYAFAGMQNNEITTTIEWNDVTLYKAFNDL 62
Db	3	LVCYFTNWAQYRQGEARPLPKDLDPISLCTHLIIYAFAGMTNHQLSTTIEWNDVTLYQGFNGL 62
Qy	63	KNRNKKLTLALAGWNFGTAPFTTWUSTSONRQTFTSVIKELROYGPDGLDLDWEYPG 122
Db	63	KQONPKLTLALAGWNFGTQKFTDMVATANNQTFVNSAIRFLRKYSGDGLDLDWEYPG 122
Qy	123	SRGSPQDKHLFTVLVKEMREAPQEAIESNRPRLMVTAAGGIGNIQAGYEIPELSKY 182
Db	123	SGGSPAVDKERFTTLVQDLANAFQOEAQTSGKERLLLSRAVPAGQYVDAGYEVDKIAGN 182
Qy	183	LDPIHMYTDLHGSGEWGTYGNSPLYKYPTETSGNAYLVNDVYVMYVKNKGAPAEKLIIVG 242
Db	183	LDVFNLMAYDFHGSWEKVYTGNSPLYKRQBSGAAASLVNDAAVQWLQKGTSPASKLIILG 242
Qy	243	PFYGHGHTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDSOE 302
Db	243	MPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYYEVCWS--KGATKQRIQDQK 300
Qy	303	VPYAYKANEWLGYNONIKSFSVKAQWLKQNNFGGAMTWAIDLDDFTGSCFQDGKFPILTSTL 362
Db	301	VPIYIFRDQWVGPDVSEFKTSYLKQKGLGGMVWALDLDLDFAGFSCNQGREYPLIQTL 360
Qy	363	NKALGI 368
Db	361	RQELSL 366
RESULT 15		
US-09-267-574-15		
; Sequence 15, Application US/09267574		
; Patent No. 6399571		
; GENERAL INFORMATION:		
; APPLICANT: Gray, Patrick W.		
; APPLICANT: Tjoelker, Larry W.		
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS		
; FILE REFERENCE: 27866/35407		
; CURRENT APPLICATION NUMBER: US/09/267,574		
; CURRENT FILING DATE: 1999-03-12		
; EARLIER APPLICATION NUMBER: 09/039,198		
; EARLIER FILING DATE: 1998-03-12		
; NUMBER OF SEQ ID NOS: 39		
; SOFTWARE: PatentIn Ver. 2.0		
; SEQ ID NO 15		
; LENGTH: 373		
; TYPE: PRT		
; ORGANISM: Homo sapiens		
US-09-267-574-15		

[illegible]

Db 243 MFTYGRSFTLASSSDTRVGPATGSGTGPFTKEGMLAYYEVCSW--KGATKQRIQDQK 300  
QY 303 VPYAYKANWLGYNIKSFSVKAOWLKONNFGGAMIWALDLDFTGSGFDOGKFPPLTSTL 362  
Db 301 VPIPRDNQWVGFDVSEFKTKVSYLKQKGLGGANWALDLDLDDFAGFSCNQGRYPLIQTL 360  
QY 363 NKALGI 368  
Db 361 RQELSL 366

Search completed: June 29, 2003, 21:03:28  
Job time : 13.7012 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 21:01:56 ; Search time 17.3846 Seconds  
(without alignments)  
2850.818 Million cell updates/sec

Title: US-10-004-219B-9  
Perfect score: 2493  
Sequence: 1 YNLICYFTNWAQYRPLGSGF.....QQHCOAGLVFTSCNCCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 424699, seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2493	100.0	452	US-10-004-219B-9	Sequence 9, Appli
2	2493	100.0	473	US-10-004-219B-4	Sequence 4, Appli
3	2089.5	83.8	455	US-10-004-219B-14	Sequence 14, Appl
4	2089.5	83.8	476	US-10-004-219B-1	Sequence 1, Appli
5	1301.5	52.2	445	US-10-004-219B-10	Sequence 10, Appl
6	987.5	39.6	383	US-10-097-340-45	Sequence 45, Appl
7	973.5	39.0	383	US-09-459-749D-17	Sequence 17, Appl
8	729.5	29.3	536	US-10-218-743-21	Sequence 21, Appl
9	729.5	29.3	555	US-10-218-743-15	Sequence 15, Appl
10	729.5	29.3	555	US-10-218-743-18	Sequence 18, Appl
11	722	29.0	490	US-10-218-743-41	Sequence 41, Appl
12	722	29.0	509	US-10-218-743-35	Sequence 35, Appl
13	722	29.0	509	US-10-218-743-38	Sequence 38, Appl
14	476	19.1	371	US-09-923-844B-2	Sequence 2, Appli
15	265.5	10.6	376	US-09-748-033-3	Sequence 3, Appli
16	203	8.1	170	US-10-218-743-44	Sequence 44, Appl
17	106	4.3	19	US-10-202-436A-13	Sequence 13, Appl
18	106	4.3	972	US-10-100-049-23	Sequence 23, Appl
19	102	4.1	1385	US-09-738-363-2	Sequence 2, Appli

20	100.5	4.0	400	12	US-10-055-430-1	Sequence 1, Appli
21	98.5	4.0	466	9	US-09-738-626-4763	Sequence 4763, Ap
22	97.5	3.9	1871	9	US-10-184-644-301	Sequence 301, App
23	97.5	3.9	1871	9	US-10-184-634-301	Sequence 301, App
24	96.5	3.9	1434	9	US-10-080-505-9	Sequence 9, Appli
25	96	3.9	237	10	US-09-924-099-9	Sequence 9, Appli
26	96	3.9	243	10	US-09-924-099-10	Sequence 10, Appl
27	95.5	3.8	299	9	US-10-007-521-12	Sequence 12, Appl
28	95.5	3.8	386	9	US-10-050-704-100	Sequence 100, App
29	95.5	3.8	695	10	US-09-945-258-18	Sequence 18, Appl
30	95.5	3.8	1157	9	US-10-184-644-391	Sequence 391, App
31	95.5	3.8	1157	9	US-10-184-634-391	Sequence 391, App
32	95	3.8	309	9	US-10-304-928-10	Sequence 10, Appl
33	95	3.8	661	10	US-09-812-238B-2	Sequence 2, Appli
34	95	3.8	661	10	US-09-862-260A-2	Sequence 2, Appli
35	95	3.8	668	9	US-10-047-539-4	Sequence 4, Appli
36	95	3.8	776	10	US-09-833-435A-4	Sequence 4, Appli
37	94.5	3.8	540	9	US-09-738-626-4450	Sequence 4450, Ap
38	91.5	3.7	425	10	US-09-888-615-92	Sequence 92, Appl
39	91.5	3.7	465	9	US-10-159-487-2	Sequence 2, Appli
40	91.5	3.7	541	9	US-10-071-485-85	Sequence 85, Appl
41	91	3.7	2024	9	US-10-123-155-259	Sequence 259, App
42	91	3.7	2105	9	US-09-428-371-3	Sequence 3, Appli
43	90.5	3.6	434	10	US-09-809-517A-23	Sequence 23, Appl
44	90.5	3.6	434	10	US-09-809-517A-26	Sequence 26, Appl
45	90.5	3.6	440	9	US-10-066-500-106	Sequence 106, App

## ALIGNMENTS

## RESULT 1

US-10-004-219B-9

; Sequence 9, Application US/10004219B

; Publication No. US20030087414A1

; GENERAL INFORMATION:

; APPLICANT: Mactozyme

; APPLICANT: Aerts, Johannes M.F.G.

; APPLICANT: Boot, Rolf G.

; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucin is involved or infection diseases

; FILE REFERENCE: 2183-5136US

; CURRENT APPLICATION NUMBER: US/10/004,219B

; CURRENT FILING DATE: 2001-11-02

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 9

; LENGTH: 452

; TYPE: PRT

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (1)..(452)

; OTHER INFORMATION: /note="Mouse AMCase"

US-10-004-219B-9

Query Match 100.0%; Score 2493; DB 9; Length 452;

Best Local Similarity 100.0%; Pred. No. 2.5e-205;

Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YNLICYFTNWAQYRPLGSGFKFPDDINPCLCTHLIYAFAGMNNETTTIENWDTLYKAFN 60

DB 1 YNLICYFTNWAQYRPLGSGFKFPDDINPCLCTHLIYAFAGMNNETTTIENWDTLYKAFN 60

QY 61 DLKKNRSKLTLLAIGGWNFGTAPFTTWVTSQNRQTFTTSVIKFLRQYGFGLDLDWEY 120

DB 61 DLKKNRSKLTLLAIGGWNFGTAPFTTWVTSQNRQTFTTSVIKFLRQYGFGLDLDWEY 120

QY 121 PGSRSGPQDKHLFTVLVKEMREAEQEAIESNRRLMWTAAVAGISNIQAGYIPELS 180

DB 121 PGSRSGPQDKHLFTVLVKEMREAEQEAIESNRRLMWTAAVAGISNIQAGYIPELS 180

QY 181 KYLDFIHVMTYDLHSGWEGYTGENSEPLYKYPTETGSNAYLVNDVYVNNYKNGAPAEKLI 240  
DB 181 KYLDFIHVMTYDLHSGWEGYTGENSEPLYKYPTETGSNAYLVNDVYVNNYKNGAPAEKLI 240  
QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSATEVWDAS 300  
DB 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSATEVWDAS 300  
QY 301 QEVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360  
DB 301 QEVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360  
QY 361 TLNKALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGSGGFCADKADGLYPVADDR 420  
DB 361 TLNKALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGSGGFCADKADGLYPVADDR 420  
QY 421 NAFWQINGITYQQHCQAGLVFDTSCNCCNP 452  
DB 421 NAFWQINGITYQQHCQAGLVFDTSCNCCNP 452

RESULT 2  
US-10-004-219B-4  
; Sequence 4, Application US/10004219B  
; Publication No. US2003008741A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 473  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase  
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence  
US-10-004-219B-4  
Query Match 100.0%; Score 2493; DB 9; Length 473;  
Best Local Similarity 100.0%; Pred. No. 2.7e-205;  
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YNLICVFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTIENNDVTLYKAFN 60  
DB 22 YNLICVFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTIENNDVTLYKAFN 81  
QY 61 DLKRNNSKLTLLAIGGNWFGTAPFTTWSTSONRQTFTTSVVKELRQYGFGLDLDWEY 120  
DB 82 DLKRNNSKLTLLAIGGNWFGTAPFTTWSTSONRQTFTTSVVKELRQYGFGLDLDWEY 141  
QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELS 180  
DB 142 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELS 201  
QY 181 KYLDFIHVMTYDLHSGWEGYTGENSEPLYKYPTETGSNAYLVNDVYVNNYKNGAPAEKLI 240  
DB 202 KYLDFIHVMTYDLHSGWEGYTGENSEPLYKYPTETGSNAYLVNDVYVNNYKNGAPAEKLI 261  
QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSATEVWDAS 300  
DB 262 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSATEVWDAS 321  
QY 301 QEVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360  
DB 322 QEVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 381

QY 361 TLNKALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGGFCADKADGLYPVADDR 420  
DB 382 TLNKALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGGFCADKADGLYPVADDR 441  
QY 421 NAFWQINGITYQQHCQAGLVFDTSCNCCNP 452  
DB 442 NAFWQINGITYQQHCQAGLVFDTSCNCCNP 473

RESULT 3  
US-10-004-219B-14  
; Sequence 14, Application US/10004219B  
; Publication No. US2003008741A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1) (455)  
; OTHER INFORMATION: /note="Human AMCase"  
US-10-004-219B-14  
Query Match 83.8%; Score 2089.5; DB 9; Length 455;  
Best Local Similarity 81.3%; Pred. No. 9.5e-171;  
Matches 369; Conservative 40; Mismatches 42; Indels 3; Gaps 1;

QY 1 YNLICVFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTIENNDVTLYKAFN 60  
DB 1 YQLTCYFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTIENNDVTLYQAFN 60  
QY 61 DLKRNNSKLTLLAIGGNWFGTAPFTTWSTSONRQTFTTSVVKELRQYGFGLDLDWEY 120  
DB 61 GLKNNKSQLKTLAIGGNWFGTAPFTTWSTSONRQTFTTSVVKELRQYGFGLDLDWEY 120  
QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELS 180  
DB 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELS 180  
QY 181 KYLDFIHVMTYDLHSGWEGYTGENSEPLYKYPTETGSNAYLVNDVYVNNYKNGAPAEKLI 240  
DB 181 QYLDYIHVMTYDLHSGWEGYTGENSEPLYKYPTETGSNAYLVNDVYVNNYKNGAPAEKLI 240  
QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSATEVWDAS 300  
DB 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSATEVWDAS 300  
QY 301 QEVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360  
DB 301 QEVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360  
QY 361 TLNKALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGGFCADKADGLYPVADDR 417  
DB 361 TLNKALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGGFCADKADGLYPVADDR 420  
QY 418 DDBNAFWQINGITYQQHCQAGLVFDTSCNCCNP 451  
DB 421 NNRNAPFWCHVNGVTYQQHCQAGLVFDTSCNCCNP 454

## RESULT 4

US-10-004-219B-1  
; Sequence 1, Application US/10004219B  
; Publication No. US2003008741A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase  
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence  
US-10-004-219B-1

Query Match 83.8%; Score 2089.5; DB 9; Length 476;  
Best Local Similarity 81.3%; Pred. No. 1e-170;  
Matches 369; Conservative 40; Mismatches 42; Indels 3; Gaps 1;  
QY 1 YNLICVFTWAQYRPGLSGFKDDINPCLCTHLIYAFAGQNNNEITTIWNDVTLKAFN 60  
DB 22 YQTCYFTWAQYRPGLSGFMEDNIDPCLCTHLIYAFAGQNNNEITTIWNDVTLKAFN 81  
QY 61 DLKRNKSLKTLAIGGWNFGTAPFTTWSTQNRQTFITSVIKFLROYGFDGLDWEY 120  
DB 82 GLKRNKSLKTLAIGGWNFGTAPFTTWSTQNRQTFITSVIKFLROYGFDGLDWEY 141  
QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGISNIQAGYEIPELS 180  
DB 142 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGISNIQAGYEIPELS 201  
QY 181 KYLDFHVMYTDLHSGWEGYTGENSPLYKYPTETGNSNAYLNVDYVMYKNGAPAEKLI 240  
DB 202 QYLDYTHVMYTDLHSGWEGYTGENSPLYKYPTETGNSNAYLNVDYVMYKNGAPAEKLI 261  
QY 241 VGPPEYGHFTILRNPDSNGIGAPTSGDGPAGATYTRQAGFWAYEICTELRSGATEVWDAS 300  
DB 262 VGPPEYGHFTILRNPDSNGIGAPTSGDGPAGATYTRQAGFWAYEICTELRSGATEVWDAS 321  
QY 301 QEVPYAYKANWGLYDNISFVSKAQLKQNNFGGAMIWALDIDFTGSCDQKFPPLTS 360  
DB 322 QEVPYAYQNVWVGYDNISFVSKAQLKQNNFGGAMIWALDIDFTGSCDQKFPPLTS 381  
QY 361 TLNKAIGISTEGCTAPDVSPSEVPTTP-----GGSGGGSGSGSGSGGFCADKADGLYPVA 417  
DB 382 TLNKAIGLASASTAPAPTEPITAPSGSGSGSGSGSGSGGFCAVRANGLYPVA 441  
QY 418 DORNAPFQINGITYOHOAGLVFDTSCNCCNW 451  
DB 442 NNRNAPFHCVNGVTYQONCOAGLVFDTSCDCCNW 475

## RESULT 5

US-10-004-219B-10  
; Sequence 10, Application US/10004219B  
; Publication No. US2003008741A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases

FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 445  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)..(445)  
; OTHER INFORMATION: /note="Human chitotriosidase"  
US-10-004-219B-10

Query Match 52.2%; Score 1301.5; DB 9; Length 445;  
Best Local Similarity 51.5%; Pred. No. 3.1e-103;  
Matches 235; Conservative 71; Mismatches 123; Indels 21; Gaps 3;  
QY 3 LICYFTNWAQYRPGLSGFKDDINPCLCTHLIYAFAGQNNNEITTIWNDVTLKAFN 62  
DB 3 LVCYFTNWAQYRQGEARFLPKDLDPSLCTHLIYAFAGTNHQLSTTWNDETLYQEFNGL 62  
QY 63 KRNKSLKTLAIGGWNFGTAPFTTWSTQNRQTFITSVIKFLROYGFDGLDWEY 122  
DB 63 KRNKSLKTLAIGGWNFGTAPFTTWSTQNRQTFITSVIKFLROYGFDGLDWEY 122  
QY 123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGISNIQAGYEIPELS 182  
DB 123 SGRSPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGISNIQAGYEIPELS 182  
QY 183 LDFHVMYTDLHSGWEGYTGENSPLYKYPTETGNSNAYLNVDYVMYKNGAPAEKLI 242  
DB 183 LDFHVMYTDLHSGWEGYTGENSPLYKYPTETGNSNAYLNVDYVMYKNGAPAEKLI 242  
QY 243 PPEYGHFTILRNPDSNGIGAPTSGDGPAGATYTRQAGFWAYEICTELRSGATEVWDAS 302  
DB 243 PPEYGHFTILRNPDSNGIGAPTSGDGPAGATYTRQAGFWAYEICTELRSGATEVWDAS 302  
QY 303 VPYAYKANWGLYDNISFVSKAQLKQNNFGGAMIWALDIDFTGSCDQKFPPLTS 362  
DB 303 VPYAYKANWGLYDNISFVSKAQLKQNNFGGAMIWALDIDFTGSCDQKFPPLTS 362  
QY 363 NKALGISTEGCTAPDV-----PSEPVTPPGSGGGSGSGSGGFCADKADGLYP 415  
DB 363 NKALGISTEGCTAPDV-----PSEPVTPPGSGGGSGSGSGGFCADKADGLYP 415  
QY 416 VADRNAPFQINGITYOHOAGLVFDTSCNCCNW 451  
DB 416 VADRNAPFQINGITYOHOAGLVFDTSCNCCNW 451  
QY 409 NPERSSFYSCAAGRLFQOQSCPTGLVFSNCKCCTW 444  
DB 409 NPERSSFYSCAAGRLFQOQSCPTGLVFSNCKCCTW 444

## RESULT 6

US-10-097-340-45  
; Sequence 45, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVAPARU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAMATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHWANDT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT





[illegible]



APPLICANT: McCall, Catherine A.  
APPLICANT: Hunter, Shirley Wu  
APPLICANT: Weber, Eric R.  
TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
TITLE OF INVENTION: AND USES THEREOF  
FILE REFERENCE: AL-2-C3  
CURRENT APPLICATION NUMBER: US/10/218,743  
CURRENT FILING DATE: 2002-08-13  
PRIOR APPLICATION NUMBER: US/09/292,225  
PRIOR FILING DATE: 1999-04-15  
PRIOR APPLICATION NUMBER: 60/098,909  
PRIOR FILING DATE: 1998-09-02  
PRIOR APPLICATION NUMBER: 60/085,295  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/098,565  
PRIOR FILING DATE: 1998-04-17  
PRIOR APPLICATION NUMBER: 09/062,013  
PRIOR FILING DATE: 1998-04-17  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 38  
LENGTH: 509  
TYPE: PRT  
ORGANISM: Dermatophagoides farinae  
US-10-218-743-38

Query Match 29.0%; Score 722; DB 9; Length 509;  
Best Local Similarity 37.7%; Pred. No. 1.7e-53;  
Matches 157; Conservative 69; Mismatches 142; Indels 48; Gaps 12;

QY 3 LICYPNNAQVRPGSLGSKFPDDINPCLCTHLYAFAGMNNETIENW-----D 52  
DB 35 IVCYVGTWSVYHK-VDPYTIEDIDPFKCTHLMYGAIDYKTIQVDFPDODNNHNSWE 93  
QY 53 VTLKAFNDLKNRNSKLTLLAIGGWNCTAPFTTWSQNRQFTITSVKELQYGF 112  
DB 94 KHGERFRNLRNKLNPDLTMTISLGGYGESEKYSMDAANPTVROQFVOSVLDLQYKFD 153  
QY 113 GLDLWEYFGSR-GSPPODKHLFTVLVKEMRAFEQAEIESNRPLMVTAAVAGISNIQ 171  
DB 154 GLDLWEYFGSRGSLGSKFPDDINPCLCTHLYAFAGMNNETIENW-----YLITAAVSPGDKID 206  
QY 172 AGYEIPELSKYLDFIHVMTYDLHGSGWYTGNSPLYKYPTETGS-NAYLNVDYVNNYWK 230  
DB 207 VAYELKELNQLFDMNMTYDHYGCGWENYVGHNAFLYKRPDETDELTYFVNNYTHYYL 266  
QY 231 NGCAEKLIVCFPGYCHTFLRNPDSNGIGAPTSQDGPAGAYTQAGFWAYEICTFLR 290  
DB 267 NNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPAGKMGPPGPGITGEEVLSYIELCOLFQ 326  
QY 291 SGATEVW----DASQEVPAVKANEMWLGVDNKSFSVKAQWLKQNNFGGAMTW 346  
DB 327 K---EWHIQYDEYNAPGYNDKIWVGYYDLASISCKLAFKLKELGVSGWMLWSLEND 383  
QY 347 TGSFCDQCKFPLTSTLANKA-----LGISTEGCTAPDVPSEPTTP 387  
DB 384 KG-HCGP-KYPL---LNKVNHNMGDEKNSYCLLGPSTTTPT-PTTSTSTTTP 433

RESULT 14  
US-09-923-844B-2  
Sequence 2, Application US/09923844B  
Patent No. US20020166143A1  
GENERAL INFORMATION:  
APPLICANT: Pioneer Hi-Bred International, Inc.  
APPLICANT: Bao, Zhongmeng  
APPLICANT: Lu, Guohua  
TITLE OF INVENTION: Sclerotinia-inducible Genes and  
TITLE OF INVENTION: Promoters and Their Uses  
FILE REFERENCE: 35718/234631  
CURRENT APPLICATION NUMBER: US/09/923,844B  
CURRENT FILING DATE: 2001-08-07  
PRIOR APPLICATION NUMBER: US 60/224,603

PRIOR FILING DATE: 2000-08-11  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 371  
TYPE: PRT  
ORGANISM: Helianthus annuus  
US-09-923-844B-2

Query Match 19.1%; Score 476; DB 9; Length 371;  
Best Local Similarity 34.1%; Pred. No. 1.3e-32;  
Matches 119; Conservative 53; Mismatches 135; Indels 42; Gaps 13;

QY 6 YFTNNAQ-YRPGLSGSKFPDDINPCLCTHLYAFAGMNNETI-----TIEWNDVTLKAFND 61  
DB 30 YWPSWAQDFLP-----PSNIQTAYFTHVYAF--LSPNNVTFFQDVHRTTASALNSFT 81  
QY 62 -LKRNNSKLTLLAIGGWNCTAP-FTTMTSTQNRQFTITSVKELQYGFGLDLWE 119  
DB 82 ALHGKNPPVKTFLSITGGGAGVKQLFSLKSSPGSRAAFIRSTIQVARNYVFDGADLWE 141  
QY 120 YPGSRGSPPODKHLFTVLVKEMRAFEQAEIESNRPLMVTAAVA-----GGISNTQA 172  
DB 142 YPETQ-----TDMNNFGLLDEWRVAVNNEATSTGKPRLLLSAATRHEPEVRDNGV----A 193  
QY 173 GYEIPELSKYLDFIHVMTYDLHGSGW-EGYTGNSPLYKYPTETGSNAYLNVDYVNNYWK 231  
DB 194 KYPVASINKLDGINAMCYDHYGWPDPATCAPAALY-----NPGSLSTSGNLQSMIS 247  
QY 232 NGAPAEKLIVCFPGYCHTFLRNPDSNGIGAPTSQDGPAGAYTQAGFWAYEICTP-LR 290  
DB 248 AGIQRQKLVMGMPYGTWTKLNPSVNGIGAPAAAGIGPG-----NEGAMLYSEVQQFNAQ 302  
QY 291 SGATEVWDASQEVPAVKANEMWLGVDNKSFSVKAQWLKQNNFGGAMTW 339  
DB 303 NNARVVYDTQTSVYSYSGTTWIGYDDVNSVORKVOYAKSLNIGSYFFW 351

RESULT 15  
US-09-748-033-3  
Sequence 3, Application US/09748033  
Patent No. US20020069431A1  
GENERAL INFORMATION:  
APPLICANT: Broadway, Roxanne M.  
APPLICANT: Gorgora, Carmenza E.  
TITLE OF INVENTION: EFFECT OF ENDOCHITINASE AND CHITOBIOSIDASE AND THEIR  
TITLE OF INVENTION: ENCODING GENES ON PLANT GROWTH AND DEVELOPMENT  
FILE REFERENCE: 19603/3091  
CURRENT APPLICATION NUMBER: US/09/748,033  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/172,003  
PRIOR FILING DATE: 1999-12-23  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 3  
LENGTH: 376  
TYPE: PRT  
ORGANISM: Streptomyces albidoflavus  
US-09-748-033-3

Query Match 10.6%; Score 265.5; DB 10; Length 376;  
Best Local Similarity 25.9%; Pred. No. 1.4e-14;  
Matches 108; Conservative 50; Mismatches 148; Indels 111; Gaps 24;

QY 6 YFTNNAQVRPGI-----GSFKPDDINPCLCTHLYAFAGMNNETITTE-----49  
DB 14 YFTENGVTYGRNYHVKNLVTSQSAEK-----ITHINYSFGNVQGGKCTIGDSFAAYDKA 66  
QY 50 -----WVDYTLKAFNDLKNRNSK-----LKTLLAIGGWNCTAPFTTMTWSQ 93  
DB 67 YTAESVDGVADTW-DQPLRGNFQNLKRLKAKYPHIKVLVWSFGGWTW-SGGFTDAV---K 121  
QY 94 NRQTFITSVIKFLRQYG-----FDGLDLWEYFGSRG-----SPPQDKHLFTVLVKEMRE 143

Db	122	NPAAFAKSHDLVEDPRWADVFDGIDLDWEYPNACGLSCDSSGPAALKN---MVOAMRA	177
Qy	144	AFQEAEIESNRPLMTAAVAGGISN---IQAGYEIPELSKYLDFTIHMVTDLHGSWEGY	200
Db	178	QFGTD-----LVTAAITADASSGGKLDAA-DYAGAAQYFDWYNVMTYDFFGAWD-K	226
Qy	201	TGENSPLYKYPTETGSGNAYLVN---DY---VMNYWKNNGAPAEKLI VGPPEYGHFTFILR	253
Db	227	TGPTAP-----HSALNSYSGIPKADPHSAAAIKAKAGVPASKLLLGIGFYGRGW---	277
Qy	254	NPSDNGIGAPTS---DGPA-GAYTRQAGFWAYYEICTFLRSGATEVWDASOEVPYAYKA	309
Db	278	-----TGVTDAPGGTATGPATGY--EAGIEDYKVLKNTCPATGTVGGTA-----YAKCG	326
Qy	310	NEWLGYDNIKSPSVKAOWLQNNFGAMIWADLDFTGSCDQGFPLTSTLNKAL	366
Db	327	SNWWSYDTPATIKTKMTWAKDQGLGGAFFW-----EFSG---DTAGGELVSAMDSGL	375

Search completed: June 29, 2003, 21:13:33  
 Job time : 18.3846 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:03:33 ; Search time 2938 Seconds  
(without alignments)  
4477.355 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLICYFTNWAQYRGLGSF.....QQHCOAGLVFDTSCNCNWP 452

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp

-Q=/cgn2\_1/USPTO\_spool/US10004219/runat\_24062003\_160228\_8159/app\_query.fasta\_1.1877  
-DB=GenEmbl -QMT=fastap -SURFIX=rg -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CGN 1 1 6828 @runat\_24062003\_160228\_8159 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*

- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.ram.\*
- 36: em.htg.rod.\*
- 37: em.htg.vrt.\*
- 38: em.sv.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2493	100.0	1530	10 AF290003	AF290003 Mus muscu
2	2488	99.8	1526	10 AF154571	AF154571 Mus muscu
3	2488	99.8	1538	10 BC011134	BC011134 Mus muscu
4	2488	99.8	1557	10 BC034548	BC034548 Mus muscu
5	2089.5	83.8	1625	9 AF290004	AF290004 Homo sapi
6	1991.5	79.9	1529	4 AB051629	AB051629 Bos tauru
7	1778.5	71.3	1369	6 AX405989	AX405989 Sequence
8	1764.5	70.8	1354	9 AB025008	AB025008 Homo sapi
9	1446	58.0	1188	9 AB025009	AB025009 Homo sapi
10	1414.5	56.7	4250	9 AK098814	AK098814 Homo sapi
11	1397	55.6	1525	10 M94584	M94584 Mus musculu
12	1379	55.3	1506	10 D87757	D87757 Mus musculu
13	1373	55.1	1209	10 AY049765	AY049765 Mus muscu
14	1371	55.0	1562	5 BJA345054	BJA345054 Bufo japo
15	1368	54.9	1209	10 AY065557	AY065557 Mus muscu
16	1301.5	52.2	1633	9 HSU29615	U29615 Human chit
17	1301.5	52.2	1636	6 AR206041	AR206041 Sequence
18	1301.5	52.2	1636	6 AR212138	AR212138 Sequence
19	1301.5	52.2	1636	6 AX108750	AX108750 Sequence
20	1301.5	52.2	1643	6 AR172107	AR172107 Sequence
21	1295.5	52.0	1656	6 AR206042	AR206042 Sequence
22	1295.5	52.0	1656	6 AR212139	AR212139 Sequence
23	1295.5	52.0	1656	6 AX108752	AX108752 Sequence
24	1286	51.6	1710	9 HSU62862	U62862 Homo sapien
25	1286	51.6	1713	6 AR172108	AR172108 Sequence
26	1253.5	50.3	93740	9 AL513202	AL513202 Human DNA
27	1200	48.1	1599	9 AK055165	AK055165 Homo sapi
28	1069.5	42.9	233175	2 AC125461	AC125461 Mus muscu
29	1051	42.2	2504	6 E09046	E09046 cDNA encodi
30	1051	42.2	2525	10 MOSOGP	D32137 Mouse mRNa
31	1043	41.8	91569	9 AL356387	AL356387 Human DNA
32	1042.5	41.8	2353	10 MAU15048	U15048 Mesocricetu
33	1042.5	41.8	2387	10 HAMOGP	D32218 Hamster mRn
34	1033.5	41.5	2366	6 E09047	E09047 cDNA encodi
35	1014	40.7	6439	3 AB074977	AB074977 Haemaphys
36	1013	40.6	2198	9 HSU09550	U09550 Human ovidu
37	1010	40.5	1994	4 BOV95BOGP	D16639 Bovine mRNa
38	1010	40.5	1994	6 E09045	E09045 cDNA encodi
39	1009.5	40.5	2228	9 BABEDOSG	M59903 Papio hamad
40	1009.5	40.5	2237	9 MU087259	U87259 Macaca mula
41	1005.5	40.3	1406	9 AF484550	AF484550 Macaca ra
42	1005.5	40.3	1909	4 AF347052	AF347052 Oryctolag
43	1003	40.2	2034	4 OAU16719	U16719 Ovis aries
44	1001	40.2	1596	4 OAU17988	U17988 Ovis aries
45	995	39.9	3946	3 AY051988	AY051988 Drosophil

ALIGNMENTS

RESULT 1





```
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
DB 1081 ATTGACCTTGATGACTCTCACTGGCTCTTCTTGATCAGGAAAATTTCTCTGACTTCT 1140
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
DB 1141 ACTTTGAACAAGCCCTTGCCATACCTCACTGAAGGTTGCACAGCTCTGACGTGCCCTCC 1200
QY 381 GluProValThrThrProProGlySerGlySerGlyGlySerGlySerGlySer 400
DB 1201 GAGCCAGTGACTACTCTCTCAGGAAGTGGAGTGGGGTGGAGCTCCCGAGAGACTCT 1260
QY 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTy-ProValAlaAspAspArg 420
DB 1261 GGAGGAGTGGATCTGTGCGCAAGCAGATGGCTCTACCTGTGGCAGATGACAGA 1320
QY 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyThrGlnGlnHisCysGlnAlaGlyLeu 440
DB 1321 AATGCTTTTGGCAGTGCATCAATGGAATCACATACAGCAGCATTTGTCAAGCAGGCTT 1380
QY 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
DB 1381 GTTTTGTATACAGCTGTAATTGCTGCAACTGGCCA 1416

RESULT 3
BC011134
LOCUS 1538 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, Similar to eosinophil chemotactic cytokine, clone
ACCESSION MGC:18771 IMAGE:4165150, mRNA, complete cds.
VERSION BC011134
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: ang@bcm.tmc.edu
Guanaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK plate: 24 Row: e Column: 11.
FEATURES
Location/Qualifiers
1..1538
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:18771 IMAGE:4165150"
/tissue type="Sallivary gland, 10 week old female mouse"
/clone lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
323..1420
/codon_start=1
```

```
/product="Similar to eosinophil chemotactic cytokine"
/protein id="AAH11134.1"
/db_xref="GI:15029822"
/translation="MVSTSQNRQTFITSVIKELRQYGFDFGLDLDWEYSGSRGSPQDK
HLFTVLVKEMREAFQEAEIESNRPLMTAAVAGISNIQAGYEIPELUSKYLDIFIHVM
TYDLHGSVEGYTGNSPLKYPTETGSNAYLNVDMYNNKNGAPAEKLVGFPEY
HTFILRNPDSNGIGAPTSQDGPAGYTRQAGFWAYEICTFLRSAGTAVWDSQVVPY
AKVANEGLYDNIKISFSVKAOLKONNFGGAMIWAIDLDLDTGSCDQKPTLTSLN
KALGISTEGCTAPPVPSEPTVTPRSGSGSGSSGSGSGSCGFCADKADGLFVADDRN
AFWOCINGITTTQQHCQAGLVDFDTCNCCNP"
BASE COUNT 393 a 382 c 365 g 398 t
ORIGIN
Alignment Scores:
Pred. No.: 5,97e-178 Length: 1538
Score: 2488.00 Matches: 451
Percent Similarity: 99.78% Conservatve: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.80% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-9 (1-452) x BC011134 (1-1538)
QY 1 TyrAsnLeuIleCysTyThrAsnTrpAlaGlnTyThrGlyLeuGlySerPhe 20
DB 62 TACAATCTGATGATGCTATTTTCAACCAATGGGCCAGTATCGCCAGGCTCGGGAGCTTC 121
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyThrAlaPheAlaGlyMet 40
DB 122 AAGCCTGATGACATTAACCCCTGCTGTGTACTACCTGATCTATGCTGCTGGGATG 181
QY 41 GlnAsnAsnGluIleThrIleGluTrpAsnAspValThrLeuTyThrAlaPheAsn 60
DB 182 CAGAACATGATGATCACCACCATAGATGATGATGATGATGATGATGATGATGATGAT 241
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 242 GACTTGAAAAACAGAACAGCAAACTGAAACCCCTCTCGCAATTTGGAGGCTGGAACCTT 301
QY 81 GlyThrAlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
DB 302 GGAACCTGCTCTCTTCACTACCATGCTTTTCCACTCTCTCAGAACCCGACACCTTCATTACC 361
QY 101 SerValIleLysPheLeuArgGlnTyThrGlyPheAspGlyLeuAspLeuAspTrpGluTy 120
DB 362 TCAGTCATCAAAATTTCTGCGTCAGTATGGGTTGATGAGCTGGACCTGGACTGGGAATAC 421
QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
DB 422 CCAGGCTCACGTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGCTCTGCTGGTGAAGAA 481
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
DB 482 ATGCGTGAAGCTTTTGGACGAGGAGCTATTGAGAGCAACAGGCCCCAGACTGATGTTACT 541
QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyThrGluIleProGluLeuSer 180
DB 542 GCTGCTGTAGCTGTGGGATTTTCCACATCCAGGCTGGCTGATGAGATCCCTGAACTTCT 601
QY 181 LysTyThrLeuAspPheIleHisValMetThrTyThrAspLeuHisGlySerTrpGluGlyTy 200
DB 602 AAGTACCTGGATTTTCATCCATGTCATGATATGATGATGATGATGATGATGATGATGAT 661
QY 201 ThrGlyGluAsnSerProLeuTyThrTyThrProThrGluThrGlySerAsnAlaTyThr 220
DB 662 ACTGGGGAGAAATAGTCTCTTTTCAAAATACCTTACTGAGACTGGTAGCAATGCTTACCTC 721
QY 221 AsnValAspTyThrValMetAsnTyThrLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
DB 722 AATGTGGATTATGTCATGAATATTGGAGAACCAATGGAGCCCGCCAGCTGAGAGCTCAT 781
QY 241 ValGlyPheProGluTyThrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
```



```
Db 782 GTTGATCCAGAGTATGACACACCTTCATCTGAGAAACCCCTCTGATATGCAATT 841
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaThrArgGlnAlaGlyPheTrp 280
Db 842 GTGCCCCCTACTCTGGTGATGGCCCTGCTGGCCCTATACACAGACAGCTGGTTCGG 901
Qy 281 AlatyTyrgluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 902 GCCTACTATGAGATTGACACCTTCTGAGAGTGGAGCCACTGAGTCTGGGATGCCTCC 961
Qy 301 GlnGluValProTyralatyrylsalalendluTrpLeuGlyTyraAspAsnlelysser 320
Db 962 CAAGAAGTGCCCTATGCTTAAAGCCAAACAGAGTGGCTTGGCTATGACAAATATCAAGAGC 1021
Qy 321 PheSerVallyAlaGlnTrpLeuLysGlnAsnAspPheGlyGlyValaMetileTrpAla 340
Db 1022 TTCAGTGTAAAGCTCAGTGGCTTAAGCAGACAAATTTGGAGGTGCCATGATCGGGCC 1081
Qy 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
Db 1082 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGAAATTTCTCTGACTTCT 1141
Qy 361 ThrLeuAsnLysAlaLeuGlyleSerThrGluGlyCysThrAlaProAspValProSer 380
Db 1142 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCCTGACGTGCTTCC 1201
Qy 381 GluProValThrProProGlySerGlySerGlyGlyGlySerSerGlyGlySerSer 400
Db 1202 GAGCCAGTGAATCTCTCCAGAGTGGAGTGGGGTGGAGCTCCCGAGAGAGCTCT 1261
Qy 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyProValAlaAspAspArg 420
Db 1262 GGAGGAGTGGATTCTGTGCCGACAAAGCAGATGGCTCTACCTGTGGCAGATGACAGA 1321
Qy 421 AsnAlaPheTrpGlnCysleAsnGlylleThrTyrglnGlnHisCysGlnAlaGlyLeu 440
Db 1322 AATGCTTTTGGCAGTGCATCAATGAATCACATACACAGCAGCATTTGTCAAGCGGCTT 1381
Qy 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
Db 1382 GTTTTGTATACAGCTGTAAATTTGCTGCAACTGGCCA 1417
RESULT 4
BC034548 1557 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, chitinase, acidic, clone MGC:19045 IMAGE:4189080,
DEFINITION mRNA, complete cdb.
ACCESSION BC034548
VERSION BC034548.1 GI:21961190
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1557)
Strausberg, R.
Direct Submission
Submitted (24-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabps-r@mail.nih.gov
cdna Library Preparation: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
```

```
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP Plate: 24 Row: h Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
Location/Qualifiers
source 1..1557
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:19045 IMAGE:4189080"
/tissue type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
4..1425
/codon_start=1
/product="chitinase, acidic"
/protein_id="AAH34548.1"
/db_xref="GI:21961191"
/db_xref="LocusID:81600"
/translation="MAKLLLVTLALLLNAQLGSAYNLICFTYNTWAQYRPGLSFKPD
DINPCLTHLYAFAGMONNEITTEWNDVLYKAFNDLKNRNSKLTLLAIGGNVFG
TAPTTVSTSONRQETITSVIKELROYGFDGLDLDWEVPGSRSPGDDKHLPTVLVK
EMRFAFEQAESENRLMVAANVAGGSIQAGVEIPELSKYLDIFHWYDVLGWSV
EGYTGNSPLKYIPETGNSVNLNVYVMYNNKNGAPAEKLIIVGFPEYGHFTILRNP
SDNGIGAPTSGDGPAGYTRQAGFWAYEICTFLRSAGTEVMDASQEPYPAKANEWL
GYNDIKSPSKAQLKONNFGGAMIWALDDFTGSCDQKGFPLTSTLNKALGISTE
GCTAPDPSPEVPTTPGSGSGSGSGSGSGSGGFCADKADGLYIPVADDRNAPMQCING
ITYQHCQAGLVFDTSCNCCWP"
BASE COUNT 410 a 382 c 366 g 399 t
ORIGIN
Alignment Scores:
Pred. No.: 6,06e-178 Length: 1557
Score: 2488.00 Matches: 451
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.80% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-9 (1-452) x BC034548 (1-1557)
Qy 1 TyrAsnLeulleCysTyPheThrAsnTrpAlaGlnTyArgProGlyLeuGlySerPhe 20
Db 67 TACAATCTGATATGCTATTTTCAACAACTGGGCCAGTATCGGCCAGGTCTGGGAGCTTC 126
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeulleTyAlaPheAlaGlyMet 40
Db 127 AAGCCTGATGACATTAAACCCCTGCTGTGATCTACCTGATCTATGCTCTTCTGGGATG 186
Qy 41 GlnAsnAsnGluileThrThrileGluTrpAsnAspValThrLeuTyLysAlaPheAsn 60
Db 187 CAGACATGATGATCACCACCATAGATGATGATGATGATGATGATGATGATGATGATG 246
Qy 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 247 GACTTGAAGAAACAGACAGCAACAACTGAAACCCCTCTGGCAATTTGGAGCTTGAAC 306
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheileThr 100
Db 307 GGAACCTGCTCTCTTTCACATACCATGCTTTTCCATCTTCTCAGAACCCGACACCTTCA 366
Qy 101 SerValIleLysPheLeuArgGlnTyrglyPheAspGlyLeuAspLeuAspTrpGluTy 120
Db 367 TCAGTCATCAAAATTTCTGGTCAGTATGGGTTTGTATGAGCTGAGCTGGAGTATAC 426
121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
```

```
Db 427 CCAGGCTCAGTGGGAGCCCTCTCAGGACAGCATCTCTCTCACTGTCTGGTGAAGGAA 486
Qy 141 MetArgGluAlaPheGluGlnGluAlaLeuSerAsnArgProArgLeuMetValThr 160
Db 487 ATGCGTGAAGCTTTGAGCAGGAGGCTATTGAGACAGCAGGCCAGACTGATGGTTACT 546
Qy 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 547 GCTGCTGTAGTGGTGGATTTCCACATCCAGGCTGGCTATGAGATCCTGAACCTTCT 606
Qy 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 607 AAGTACTCGATTCATCCATGTCATGACATATGACCTCATGGCTCTCGGAGGGCTAC 666
Qy 201 ThrGluGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 667 ACTGGGAGAGATAGTCTCTTTACAATACCTCTGAGACTGGTAGCAATGCTTACCTC 726
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
Db 727 AATGTGGATTATGTATGTAACATATTGGAAGAACAAATGGAGCCCGAGCTGAGAAGCTCAT 786
Qy 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
Db 787 GTTGGATTCCAGAGTATGACACACCTTTCATCTCGAAGAACCCCTCTGATATGGAATT 846
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
Db 847 GGTGCCCTTACCTCTGTGTATGCGCTCTGTGGCCCTATACACAGACAGCTGGTTCGG 906
Qy 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 907 GCTACTATGAGATTGACCTTCTTGAGAGTGGAGCCACTGAGGTCTGGGATGCTCTCC 966
Qy 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
Db 967 CAAGAAGTCCCTATGCTTATAGGCAAGCAAGTGGCTTGGCTATGACAAATATCAAGAGC 1026
Qy 321 PheSerValIleAlaGlnTrpLeuLysGlnAsnAsnPheGlyAlaMetIleTrpAla 340
Db 1027 TTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCC 1086
Qy 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
Db 1087 ATTGACCTTGTATGACTTCACTGGCTCTTCTGTGATCAGGGAATTTCTCTGACTTCT 1146
Qy 361 ThrLeuAsnIleAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
Db 1147 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCCTGACGTGCTTCC 1206
Qy 381 GluProValThrProProGlySerGlySerGlyGlyGlySerSerGlyGlySerSer 400
Db 1207 GAGCCAGTGACTACTCTCTCAGGAAGTGGAGTGGGGGTGGAGCTCCGAGGAAGCTCT 1266
Qy 401 GlyGlySerGlyPheCysAlaAlaPheLysAlaAspGlyLeuTyrProValAlaAspAspArg 420
Db 1267 CGAGGCAGTGGATTCTGTGCCGACAAAGCAGATGGCTCTTACCTGTGGCAGATGACAGA 1326
Qy 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440
Db 1327 AATGCTTTTGGCAGTGCATCAATGGAATACATACACAGCAGCATTTGTCAAGCAGGCTT 1386
Qy 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
Db 1387 GTTTTGTATACCACTGTAAATTTGCTGCAACTGGCCA 1422
RESULT 5
AF290004
LOCUS
DEFINITION Homo sapiens acidic mammalian chitinase precursor, mRNA, complete
cgs
ACCESSION AF290004
VERSION AF290004.1 GI:12597292
```

```
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1625)
AUTHORS Boot,R.G., Blommaert,E.F., Swart,E., Ghauharali-van der Vlugt,K.,
Bijl,N., Moe,C., Place,A. and Aerts,J.M.
TITLE Identification of a novel acidic mammalian chitinase distinct from
chitotriosidase
JOURNAL J. Biol. Chem. 276 (9), 6770-6778 (2001)
MEDLINE 21125893
PUBMED 11085997
REFERENCE 2 (bases 1 to 1625)
AUTHORS Boot,R.G., Verhoek,M., Swart,E. and Aerts,J.M.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2000) Dept. of Biochemistry, Academic Medical
Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105
AZ, The Netherlands
FEATURES
source
1..1625
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="stomach; lung"
104..1534
/EC_number="3.2.1.14"
/note="chitinase precursor; AMCase; chitinase activity
demonstrated by expression in COS cells; similar to
sequences presented in GenBank Accession Numbers AB025008
and AB025009"
/codon_start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60019.1"
/db_xref="GI:12597293"
/translation="MTKLILTLGLVLIINLQLSAYQLTCTFTWAQYRPGLRMPD
NIDPLCTLLIYAFAGRONNEITTEWDVLYQAFNGLNKNKSOLKTLAIGWNFG
TAPFTAMVSTPENQFTITSVKFLROYDFDLDFWEYPGSRGSPQDKHLFTLVQ
EMRFAEQEAKIQINKPRLMTAAVAGISNIQSGVEIPLSQYLDYIHMTYDLHGSW
EGYTGENSPKYPTDTSNAYLVNVDYVMYKONGAPAEKLIYGFPTYGHNFLSNP
SNTGKISDTSAGAGPAGPVAKESGIWAYVEICTFLKNGATCGMDAPQEPVYAYQNVV
GYDNKISFDIKAWLKHNFQGMVWALDDFTGTGTCNQKFPILITLKALGLQSA
SCTPAQPIETPTAASGSGNGSGSSGSGSGGSCFCAVRANGLYPVANNRNFVHC
VNGVTYQNCQAGLVFDFTSCDCNWA"
104..166
sig_peptide 387 a 468 c 399 g 371 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 5,46e-148 Length: 1625
Score: 2089.50 Matches: 369
Percent Similarity: 90.0% Conservative: 40
Best Local Similarity: 81.2% Mismatches: 42
Query Match: 83.8% Indels: 3
DB: 9 Gaps: 1
US-10-004-219B-9 (1-452) x AF290004 (1-1625)
Qy 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 167 TACCAGCTGACATGCTACTTACCACACTGGGCCAGTACCAGCCAGGCTGGGGCGCTTC 226
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 227 ATGCTCTGACACATCGACCCCTGCTCTGTACCCACCTGATCAGCTTGTGCGGAGG 286
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 287 CAGAACACAGAGATCACCACTCGAATGGAACGATGTGACTCTCTACCAAGCTTTCAT 346
Qy 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 347 GGCCTGAAAAATAGACAGCCAGCTGAAACTCTCTCTGGCCATTTGGAGGCTGGAACTTC 406
```

81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100  
 407 GGGACTGGCCCTTTCACTGCCATGTTTCTACTCTGAGAACCGCCAGACTTTTCATCACC 466  
 101 SerValIleLysPheLeuArgGlnThrGlyPheAspGlyLeuAspLeuAspTyrGluTyr 120  
 467 TCAGTCATCAAAATCTCGCCAGTATGAGTTTGAAGGCTGGACTTTGACTGGGAGTAC 526  
 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140  
 527 CTTGGCTCTGTTGGAGGCTCTCTGAGACAGCATCTTCTACTGTCTGTTGGAGGAG 586  
 141 MetArgGluAlaPheGluGlnGluAlaLeuSerAsnArgProArgLeuMetValThr 160  
 587 ATGCGTGAAGCTTTTGGAGGAGGAGCAAGCAATCAAGCCAGGCTGATGCTACT 646  
 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
 647 GCTGAGTAGTCTGCTGGCATCTTCAATATCCAGTCTGCTATGAGATCCCCCACTGTCA 706  
 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyr 200  
 707 CAGTACCTGGACTCATCTATGCTATGACTTACGCTCATGCTCTCTGGAGGAGCTAC 766  
 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220  
 767 ACTGGAGAGAACAGCCCTCTTACAAATACCCGAGTGTACACCCGAGCAAGCCCTACCTC 826  
 221 AsnValAspTyrValMetAsnTyrTyrLysAsnAsnGlyAlaProAlaGluLysLeuIle 240  
 827 AATGTGGATTATGTATGAATCTACTTGGAGGAGCAATGGAGCACCACTGAGAAGCTCATC 886  
 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260  
 887 GTTGATTTCCCTTACCTATGAGACAACTTCACTCTGAGCAACCCCTTCCACACTGGAAT 946  
 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTyr 280  
 947 GGTGCCCCCACCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1006  
 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTyrAspAlaSer 300  
 1007 GCTTACTACGAGATCTGTACTTCTGAAATAATGGAGCCACTCAGGGATGGATGCCCT 1066  
 301 GlnGluValProTyrAlaTyrLysAlaGlnTyrLeuGlyTyrAspAsnIleLysSer 320  
 1067 CAGGAAGTCCCTTATGCTTATCAGGCAATGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1126  
 321 PheSerValLysAlaGlnThrLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTyrAla 340  
 1127 TTCGATATTAAAGCTCAATGGCTTAAGCACACAAATTTGGAGGCGCCATGCTGGGCC 1186  
 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyPheProLeuThrSer 360  
 1187 ATTGATCTGGATGACTTCACTGCGCACTTCTGCAACAGGCAAGTTTCCCTTAATCTCC 1246  
 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380  
 1247 ACCCTGAAGAGGCGCTTCCGCTGAGAGTCAAGTGTGACGCGCTCCAGCTCAGCCCAT 1306  
 381 GluProValThrThrProPro-----GlySerGlySerGlyGlySerGly 397  
 1307 GAGCCATAAATCTGCTCTCCAGTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1366  
 398 GlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAla 417  
 1367 GGCAGCTCGGAGGAGGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1426  
 418 AspAspArgAsnAlaPheThrGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGln 437  
 1427 AATACAGAAATGCTTCTGCGCACTGCGTGAATGGAGTCACTGCTGCTGCTGCTGCTGCT 1486  
 438 AlaGlyLeuValPheAspThrSerCysAsnCysAsnTyr 451

DB 1487 GCCGGCTGTCTTCGACACAGCTGTGATTGCTGCAACTGG 1528  
 RESULT 6  
 LOCUS AB051629 1529 bp mRNA linear MAM 07-NOV-2001  
 DEFINITION Bos taurus cbp b04 mRNA for chitin binding protein b04, complete cds.  
 ACCESSION AB051629 GI:16754894  
 VERSION AB051629.1  
 KEYWORDS Bos taurus liver cDNA to mRNA.  
 SOURCE Bos taurus  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.  
 REFERENCE 1 Suzuki, M., Morimatsu, M., Yamashita, T., Iwanaga, T. and Syuto, B. A novel serum chitinase that is expressed in bovine liver FEBS Lett. 506 (2), 127-130 (2001)  
 MEDLINE 21475601  
 REFERENCE 2 (bases 1 to 1529)  
 AUTHORS Suzuki, M., Morimatsu, M. and Syuto, B.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-NOV-2000) Masako Suzuki, Iwate University, Department of Veterinary Medicine, Faculty of Agriculture, Ueda 3-18-8, Morioka, Iwate 020-8550, Japan (E-mail: u9998001@iwate-u.ac.jp, Tel: 81-19-621-6212)  
 FEATURES  
 source 1. 1529 /organism="Bos taurus" /db\_xref="taxon:9913" /tissue\_type="liver" 1. 1529 /gene="cbp b04" 13. 1431 /gene="cbp b04" /codon\_start=1 /product="chitin binding protein b04" /protein\_id="BAB71805.1" /db\_xref="GI:16754895" /translation="MAKLIFLTGLAFLLNAQLGSAYQLVCFNSNWAQYRPGLSKPKPD NIDPLCTHLIYAFAGMSNSETITENWDVALYSFNDLKKNSQLKLLAIGGNFVQ TAPFTAMVTPENRKTFTSSVIFKLHGYDGLDFDMEYPGFRGSPQDKHLFTVLVQ ETRAFGEAKQTNKPLRLTAAVAGISNIQAGVEIPOLSOYLDPIHMTVDPHGWSV EGYTGNSPLKYPTDTSNTYLVNVEYAMNWKNGAPAEKLIIGFPAYGHNFLIRDA SNNGIAPTSKAGGAPGPTREAGFWAYTEICAFKLDGATEAWDDSQNVPIYAKGTETWY GYDYNVPRIKAOIWKNNFGAMVWAIIDLDFTGTCFCNGKFPPLINTLKDALGLKSA TCNASTOSSEPNSSPNSGSGNSKSSSEGRGYCAGKADGLYPVADNRNAPMNCVNGI TYKQNCUTGLVFDTSCHCNWA" 1508. 1513 /gene="cbp b04"  
 polyA\_signal  
 BASE COUNT 391 a 405 c 365 g 368 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,17e-140 Length: 1529  
 Score: 1991.50 Matches: 348  
 Percent Similarity: 87.14% Conservative: 45  
 Best Local Similarity: 77.16% Mismatches: 57  
 Query Match: 79.88% Indels: 1  
 DB: 4 Gaps: 1  
 US-10-004-219B-9 (1-452) x AB051629 (1-1529)  
 QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20  
 DB 76 TACCAGCTGGTATGCTACTTCTTAACCTGGGCCAGTACCGCTCCAGGCTGGGAAGCTTC 135  
 QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
 DB 136 AAGCCTGCAACATCGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 195



```

QY 139 LysGluMetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMet 158
Db 307 CAGGAATCGGTGAAGCTTTTTCAGCAGGAGGCAAGCAGATCAACAGCCAGGCTGATG 366
QY 159 ValThrAlaAlaValAlaGlyClyIleSerAsnIleGlnAlaGlyTyrGluIleProGlu 178
Db 367 GTCACTGCTGCAGTAGCTGCTGGCATCTCCATATCCAGTCTGGCTATGAGATCCCCAA 426
QY 179 LeuSerLysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTyrGlu 198
Db 427 CTGTCACTGACTGACTACATCTCATCTGATGACTAGCAGCTTCATGGCTCTCTGGAG 486
QY 199 GlyTyrThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAla 218
Db 487 GGTACTACTGGAGAGAACAGCCCTCTTACAAATACCCGACTGACACCGCGCAGCAACGCC 546
QY 219 TyrLeuAsnValAspTyrValMetAsnTyrTyrLysAsnAsnGlyAlaProAlaGluLys 238
Db 547 TACCTCAATGTGATATTATGTCTAAGTAAGTACTGCAAGGACATGGAGCACCGACTGAG 606
QY 239 LeuIleValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsn 258
Db 607 CTCATCGTTGGATTCCTTACCTATGACACAACTTCATCTGAGCAACCCCTCCAACT 666
QY 259 GlyIleGlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGly 278
Db 667 GGAATTTGGTGGCCCCCCTCTGGTCTGCTGCTGGCCCTATGCCAAGAGTCTGGG 726
QY 279 PheThrAlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTyrAsp 298
Db 727 ATCTGGGCTTACTACGAGATCTGTACTCTTCTGGAATAATGGAGCCACTCAGCGATGG 786
QY 299 AlaSerGlnGluValProTyrAlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIle 318
Db 787 GCCCTCAGGAAGTGCCTTATGCCATATGAGGCAATGTGGGTGGCTATGACAACTC 846
QY 319 LysSerPheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyClyAlaMetIle 338
Db 847 AAGAGCTTCGATATTAAAGCTCAATGGCTTAAGCAACAATAATTTGGAGGCCATGGTC 906
QY 339 TrpAlaIleAspLeuAspPheThrGlySerPheCysAspGlnGlyLysPheProLeu 358
Db 907 TGGGCCATTTGATCTGATGACTTCACTGCACTTTCTGCAACCGGCGCAAGTTTCCCTTA 966
QY 359 ThrSerThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal 378
Db 967 ATCTCACCTCGAAGAGGCCCTCGGCCCTGACAGTGCAGTTCAGCGCTCCAGCTCAG 1026
QY 379 ProSerGluProValThrThrProProProProProProProProProProProProPro 395
Db 1027 CCATTGAGCCCAATAACTGCTGCTCCAGTGCAGCGGCAACGGGAGCGGGAGTAGCAGC 1086
QY 396 SerGlyGlySerSerGlyCysGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
Db 1087 TCTGAGGCGAGCTCGGAGGCGAGTGTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
Db 1147 GTGGCAATAACAGAAATGCCCTCTGGCACTCGGCAATGGAGTACAGTACCGAGCAGAAC 1206
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
Db 1207 TGCCAGGCGGCGGCTGCTCTTCGACACCGAGCTGTGATGTGCTGCACTGG 1254

```

RESULT 8  
 AB025008  
 LOCUS  
 DEFINITION  
 Homo sapiens TSA1902-L mRNA for novel member of chitinase family,  
 complete cds.  
 ACCESSION  
 AB025008  
 VERSION  
 AB025008.1 GI:6467176  
 KEYWORDS  
 novel member of chitinase family; TSA1902-L.  
 SOURCE  
 Homo sapiens cDNA to mRNA.

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Saito,A., Ozaki,K., Fujiwara,T., Nakamura,Y. and Tanigami,A.
JOURNAL Isolation and mapping of a human lung-specific gene, TSA1902,
MEDLINE encoding a novel chitinase family member
AUTHORS Gene 239 (2), 325-331 (1999)
TITLE Direct Submission
JOURNAL Saito,A., Ozaki,K., Fujiwara,T., Takahashi,E. and Tanigami,A.
REFERENCE 20018184
AUTHORS 2 (bases 1 to 1354)
TITLE Submitted (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co.,
JOURNAL Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho,
Tokushima, Tokushima 771-0192, Japan (E-mail:saito@otsuka.gr.jp,
Tel:81-88-665-2888, Fax:81-88-637-1035)
FEATURES
source
1..1354
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1..1354
/gene="TSA1902-L"
CDS
157..1263
/gene="TSA1902-L"
/codon_start=1
/product="novel member of chitinase family"
/protein_id="BAA06980.1"
/db_xref="GI:6467177"
/translation="MVSPTENRQTFITSVIKFLRQYEDGLDFDWEYFGSRGSPQDDK
HLFTVLQVEMREAFQEKQINQKRLMTVAAGISNIQSGYELIPQISQYLDYIHVM
TYDLHGWEYGTGNSPLKYPTDTGNSAYLNVYVWYKNDGAPAKLIVGFPTYG
HNFILSNPNTGIGATSGAGPAGPYAKSGIWAYYEICTFLKNGATQGDAPQVPY
AYQGNVWGYDYNKVSFDIKAWLKHNRKFGAMVWALDLDFTGTFNCGKPLITLK
KALGLQASACTAFAPQLEPITAAPSGSGSGSSSSSGSSGSGGFCAGRANGLYPVAN
NRNFAHCVNGVTYQQNQAGLVFDTSCDCCNWA"
BASE COUNT 311 a 387 c 348 g 308 t
ORIGIN
Alignment Scores:
Pred. No.: 1,14e-123 Length: 1354
Score: 1764.50 Matches: 312
Percent Similarity: 89.72% Conservative: 37
Best Local Similarity: 80.21% Mismatches: 37
Query Match: 70.78% Indels: 3
DB: 9 Gaps: 1
US-10-004-219b-9 (1-452) x AB025008 (1-1354)
QY 66 AsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPhe 85
Db 91 AACAGCCAGCTGAAACCTCTCTGGCCATTGGAGCTGGAACCTTCAGGACTGCCCTTTC 150
QY 86 ThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerValIleLysPhe 105
Db 151 ACTGCGCATGGTTTCTACTCTCTGAGAACCGCCAGACTTTTCATCACCTCAGTCATAAATTC 210
QY 106 LeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTyrGluTyrProGlySerArgGly 125
Db 211 CTGCCCATGATGATGATTTTGGAGCGGCTGGAGCTTTGAGTGGAGTACCCCTGCTCTCTG 270
QY 126 SerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGluAlaPhe 145
Db 271 AGCCCTCTCAGGACACAGCATCTCTTCACTGCTGCTGGTGCAGGAATCGGTGAAGCTTTT 330
QY 146 GluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGly 165
Db 331 GAGCAGGAGGCGCAAGCAGATCAACAAAGCCAGGCTGATGGTCACTGCTGCAGTAGCTGT 390
QY 166 GlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPhe 185
Db 391 GGCATCTCAATATCCAGTCTGGCTATGAGATGCCCAACTGTTCACGTACCTGACTGACTAC 450
QY 186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205

```

```

Db 451 ATCCATGTTCATGACCTACGACCTCCATGCTCTCTGGAGGCTTACACTGGAGAACAGC 510
Qy 206 ProLeuTyrLysThrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225
Db 511 CCCCTCTACAAATACCCGACTGACACCGGACAGCCCTACCTCAATGTGATATATGTC 570
Qy 226 MetAsnTyrTyrLysAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGlu 245
Db 571 ATGAACCTACTGGAGGACATGGAGCACAGCTGAGAGCTATCTGTGGATCCCTTACC 630
Qy 246 TyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSer 265
Db 631 TATGGACACAACTTCATCTGAGCAACCCCTCCACACTGGAATGGTCCGCCACCTCT 690
Qy 266 GlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheThrAlaTyrGluIle 285
Db 691 GGTGCTGGTCTGCTGGGCCCTATGCAAGGAGTCTGGGATCTGGGCTTACTACGAGATC 750
Qy 286 CysThrPheLeuArgSerGlyAlaThrGluValTyrAspAlaSerGlnGluValProTyr 305
Db 751 TGTACCTTCTGAATAATGGACCACTCAGGATGGGATGCCCTCAGGAAGTGCCTTAT 810
Qy 306 AlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
Db 811 GCTATCAGGCAATGTGTGGTGGCTATGCAACGCTCAAGAGCTTCGATATTAAAGCT 870
Qy 326 GlnTyrLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTyrAlaIleAspLeuAsp 345
Db 871 CAATGGCTTAAGCACACAAATTTGAGAGCGCATGCTCTGGGCCATTCATCTGGATGAC 930
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
Db 931 TTCACCTGGCACTTCTGCAACCGGCAAGTTCCTTAATCTCCACCTCGAAGAGGCC 990
Qy 366 LeuGlyIleSerThrGluGlyCysThrAlaProAspValProSerGluProValThrThr 385
Db 991 CTTGGCTGCAGAGTCAAGTTCACCGCTCCAGCTCAGCCCATTCAGGCAATACTGCT 1050
Qy 386 ProPro-----GlySerGlySerGlyGlyGlySerSerGlyCysSerGlyGly 402
Db 1051 GTCCCACTGGCAGCGGAGGAGCGGAGTAGCAGCTCTGGAGCGAGCTCGGGAGGC 1110
Qy 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAla 422
Db 1111 AGTGGATTCTGTGGCAGACCAAGCGCTTACCCCTGGCAATTAACAGAAATGCC 1170
Qy 423 PheTyrGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442
Db 1171 TTCCTGCACTGCGTGAATGGAGTCAGCTACAGCAGAACTGCCAGCGCGGCTTGTCTTC 1230
Qy 443 AspThrSerCysAsnCysCysAsnTyr 451
Db 1231 GACACAGCTGTGATTGTGCTCAACTGG 1257

RESULT 9
AB025009
LOCUS
DEFINITION Homo sapiens TSA1902-S mRNA for novel member of chitinase family, complete cds.
ACCESSION AB025009
VERSION AB025009.1 GI:6467178
KEYWORDS novel member of chitinase family; TSA1902-S.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1188)
AUTHORS Saito,A., Ozaki,K., Fujiwara,T., Nakamura,Y. and Tanigami,A.
TITLE Isolation and mapping of a human lung-specific gene, TSA1902, encoding a novel chitinase family member
JOURNAL Gene, 239 (2), 325-331 (1999)
MEDLINE 20018184

```

```

REFERENCE 2 (bases 1 to 1188)
AUTHORS Saito,A., Ozaki,K., Fujiwara,T., Takahashi,E. and Tanigami,A.
TITLE Direct Submission
JOURNAL Submitted (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co., Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-0192, Japan (E-mail:saito@otsuka.gr.jp, Tel:81-88-665-2888, Fax:81-88-637-1035)
FEATURES
source
1..1188
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1..1188
/gene="TSA1902-S"
CDS
150..1097
/gene="TSA1902-S"
/codon_start=1
/product="novel member of chitinase family"
/protein_id="BAA86981.1"
/db_xref="GI:6467178"
/translat="MREAFEOEAKQIKPRLMTVAAVAAGISNIQSGYEIPOLSOYLD
YIHWYDLHGSWEGYTGENSEPLKYPTDTSNAYLVNVDYVMYWKDNGAPAEKLVG
PFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYVEICTFLKNGATQGDAP
QEVPIAYGVNVDNYSFDIKQMLKHNKFGGMWALDLDFTTCTFCNGKRPFL
ISTLKALGLQSASCTAPQIEPIITAPSGSGSGSSSGSGSGSGSGSGSGSGSGSG
YFVANNRNFVHCVNGVTYQNCQAGLVFDTSCDCNNA"
BASE COUNT 280 a 337 c 309 g 262 t
ORIGIN
Alignment Scores:
Pred. No.: 8,286-100 Length: 1188
Score: 1446.00 Matches: 262
Percent Similarity: 76.35% Conservative: 35
Best Local Similarity: 67.35% Mismatches: 34
Query Match: 58.00% Indels: 59
DB: 9 Gaps: 2
US-10-004-219b-9 (1-452) x AB025009 (1-1188)
Qy 66 AsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPhe 85
Db 91 AACACCAAGCTGAAACCTCTCTGGCCATTGGAGGCTGGAACTTCAGGACTGC-CCC--- 146
Qy 86 ThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerValIleLysPhe 105
Db 146 ----- 146
Qy 106 LeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTyrGluTyrProGlySerArgGly 125
Db 146 ----- 146
Qy 126 SerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGluAlaPhe 145
Db 147 -----GAAATGGTGAAGCTTTT 164
Qy 146 GluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGly 165
Db 165 GAGCAGGAGGCGCAAGCAGATCAACAAGCCAGGCTGATGGTCACTGCTGCAGTAGCTGCT 224
Qy 166 GlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPhe 185
Db 225 GGCATCTCCAATATCCAGCTGGCTATGAGATCCCACTGTCACAGTACCTGACTAC 284
Qy 186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205
Db 285 ATCCATGTCATGACCTACGACCTCCATGGCTCTGGAGGGCTACACTGGAGAGAACAGC 344
Qy 206 ProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225
Db 345 CCCCTCTCAAAATACCCGACTGACACCGGACAGCCCTACCTCAATGTGATATATGTC 404
Qy 226 MetAsnTyrTyrLysAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGlu 245
Db 405 ATGAACTACTGGAGGCAATGGAGCACAGCTGAGAAGCTCATCGTTGGATTCCCTACC 464

```

```
Qy 246 TyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSer 265
Db 465 TATGGACACAACTTCTATCTGAGCAACCCCTCCACACTGGAAATGGTGGCCCACTCT 524
Qy 266 GlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrGluIle 285
Db 525 GGTGTGTGCTCTGCTGGCCCTATGCAAGAGTCTGGGATCTGGGCTTACTACGAGATC 584
Qy 286 CysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyr 305
Db 585 TGTACTCTCTGAAATATGAGCCACTCAGGATGGATGGCCCTCAGGAATGGCTTAT 644
Qy 306 AlaTyrLysAlaAsnGlnTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
Db 645 GCCTATCAGGGCAATGCTGTGGGTGGCTATGACCAAGTCAAGTCAAGAGCTTCGATTAAGGCT 704
Qy 326 GlnTrpLeuLysGlnAsnAsnPhedGlyAlaMetIleTrpAlaIleAspLeuAsp 345
Db 705 CAATGGCTTAAGCACAACAAAATTTGGAGCGCATGGTCTGGGCCATTGATCTGGATGAC 764
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
Db 765 TTCATGGGCATCTTCTGCNACAGGCGCAAGTTCCCTTAATCTCCACCTCGAAGAGGCC 824
Qy 366 LeuGlyIleSerThrGluGlyCysThrAlaProAspValProSerGluProValThrThr 385
Db 825 CTTGGCTGCAGAGTGCAGATTGTCAGCGCTCAGCTCAGCCCATTTAGCCATAACTGCT 884
Qy 386 ProPro-----GlySerGlySerGlyGlySerGlyGlySerGlySerGlyGly 402
Db 885 GCTCCAGTGGCAGCGGGAACCGGAGCGGAGTAGCAGCTCTGGAGCGCAGCTCGGGAGGC 944
Qy 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspArgAsnAla 422
Db 945 AGTGGATCTGTGGTGGCAGACCAAGCGCTCTACCCGCGTGCAGAAATAACAGAAATGCC 1004
Qy 423 PheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442
Db 1005 TTCCTGGCACTGCTGNATGAGTCACTACAGCAGCAGACTGCGAGCGCGGCTTCTTC 1064
Qy 443 AspThrSerCysAsnCysCysAsnTrp 451
Db 1065 GACACAGCTGTGATGCTGCAACTGG 1091

RESULT 10
AK098814
LOCUS
DEFINITION
Homo sapiens cDNA FLJ25948 fis, clone STM04207, highly similar to
Homo sapiens acidic mammalian chitinase precursor.
ACCESSION
AK098814
VERSION
KEYWORDS
SOURCE
Homo sapiens
clone:STM04207.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M.,
Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and
Sugano, S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 4250)
REFERENCE
Sugano, S. and Suzuki, Y.
Direct Submission
Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
```

```
Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.
FEATURES
Location/Qualifiers
1..4250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="STM04207"
/tissue_type="stomach mucosa"
/clone_lib="STM"
/note="cloning vector: pME18SFL3"
BASE COUNT 1262 a 868 c 966 g 1154 t
ORIGIN
Alignment Scores:
Pred. No.: 8,27e-97 Length: 4250
Score: 1414.50 Matches: 249
Percent Similarity: 87.69% Conservative: 36
Best Local Similarity: 76.62% Mismatches: 37
Query Match: 56.74% Indels: 3
DB: 9 Gaps: 1
US-10-004-219B-9 (1-452) x AK098814 (1-4250)
Qy 130 AsplysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAla 149
Db 3174 GATGTCCTATTATGCTTATTATCTGTAGGAATGCGTGAAGCTTTTGAGCAGAGGCC 3233
Qy 150 IleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsn 169
Db 3234 AAGCAGATCAACAAGCCCGCAGCTGATGCTCACTGCTGCAGTAGCTGCTGGCATCTCCAAT 3293
Qy 170 IleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisValMet 189
Db 3294 ATCCAGTCTGGCTATGAGATCCCCCACTGCACAGTACCTGGAGTACATCATCTGTCATG 3353
Qy 190 ThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209
Db 3354 ACCTACGACCTCCATGGCTCTCTGGAGGGCTACACTGGAGAGAACAGCCCCCTCTACAAA 3413
Qy 210 TyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
Db 3414 TACCCGACTACACCGCGCAGCAACGCCCTACCTCAATGTGGATTATGTCAATGAACCTACTGG 3473
Qy 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThr 249
Db 3474 AAGGCAATGGAGCACCAGCTGAGAAGCTCATCGTTGGATTCCCTACCTATGGACACAAAC 3533
Qy 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
Db 3534 TTCATCTGAGCAACCCCTCCCAACTGGAATTTGGTGGCCCACTCTGGTCTGGTCTCT 3593
Qy 270 AlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluIleCysThrPheLeu 289
Db 3594 GCTGGGCCCTATGCAAGAGTCTGGGCTTACTACGAGATCTGTACCTTCTCTG 3653
Qy 290 ArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyrAlaTyrLysAla 309
Db 3654 AAAATGGAGCCACTCAGGATGGGATGCCCTCAGGAAGTGCCTTATGCCCTATCAGGCGC 3713
Qy 310 AsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAlaGlnTrpLeuLys 329
Db 3714 AATGTGGTGGTGGCTATGACCAACATCAAGAGCTTCGATATTAGGCTCAATGGCTTTAG 3773
Qy 330 GlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspPheThrGlySer 349
Db 330 GlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspPheThrGlySer 349
```

COMMENT

FEATURES  
source



```
Db 3774 CACAACAAATTTGGAGCGCCATGTCGTGGCCCATGATCTGGATGACTTCACTGGCACT 3833
Qy 350 PheCysAspGlnGlySerPheProLeuThrSerThrLeuAsnLysAlaLeuGlyIleSer 369
Db 3834 TTCTGACACAGCGCAAGTTTCCTCCCTAACTCCACCTGAAGAGCCCTCGGCCCTGCAG 3893
Qy 370 ThrGluGlyCysThrAlaProAspValProSerGluProValThrThrProPro----- 387
Db 3894 AGTGCAAGTTGACAGCGCTCAGCTCAGCCATTGAGCCCAATACTGCTGCCAGTGCG 3953
Qy 388 ---GlySerGlySerGlyGlySerSerGlyGlySerSerGlyGlySerGlyPheCys 406
Db 3954 AGCGGAACGGGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGGAGCGAGTGGATTCTGT 4013
Qy 407 AlaAspLysAlaAspGlyLeuThrProValAlaAspAspArgAsnAlaPheTrpGlnCys 426
Db 4014 GCTGTGAGAGCAACCGCTCTACCCCGTGGCAATAAAGAGAAATGCTTCTGGCACTGC 4073
Qy 427 IleAsnGlyIleThrTrpGlnGlnHisCysGlnAlaGlyLeuValPheAspThrSerCys 446
Db 4074 GTGAATGAGTACAGTACCGACGAGCACTGCCAGCGCGGCTTGTCTTCGACACAGCTGT 4133
Qy 447 AsnCysCysAsnTrp 451
Db 4134 GATTGCTGCAACTGG 4148

RESULT 11
M94584
LOCUS M94584 1525 bp mRNA linear ROD 15-JUN-2001
DEFINITION Mus musculus secretory protein precursor (Yml) mRNA, complete cds.
ACCESSION M94584
VERSION M94584.2 GI:11140876
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 1525)
AUTHORS Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
TITLE Molecular characterization of a secretory protein (YM-1) transiently expressed by activated murine peritoneal macrophages
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1525)
AUTHORS Chang,N.C., Hung,S.I., Hwa,K.Y., Kato,I., Chen,J.E., Liu,C.H. and Chang,A.C.
TITLE A macrophage protein, Yml, transiently expressed during inflammation is a novel mammalian lectin
JOURNAL J. Biol. Chem. 276 (20), 17497-17506 (2001)
MEDLINE 21264517
PUBMED 11297523
REFERENCE 3 (bases 1 to 1525)
AUTHORS Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1993) Microbiology and Immunology, National Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan 112, R.O.C.
REFERENCE 4 (bases 1 to 1525)
AUTHORS Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2000) Microbiology and Immunology, National Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan 112, R.O.C.
REMARK Sequence update by submitter.
COMMENT On Nov 13, 2000 this sequence version replaced gi:202441.
FEATURES
source
location/Qualifiers
1..1525
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="PEC in lambda gtl1 and ZAPII"
/dev_stage="adult"
1..1525
/gene="Yml"
13..1209
CDS
```

```
/gene="Yml"
/codon_start=1
/product="secretory protein precursor"
/protein_id="AAB62394.2"
/db_xref="GI:11140877"
/translation="MAKILVTGLAILNLVOLGSSYQLMCMYTSWAKDRPIEGSPKPG
NIDFLCTHLIYAFAGMNEIYTHQDLRYEALNGKDKNTLKTLLAIGWKFQ
PAPFSAMVSTPQNRQIFIQSVIRFLQYDFDLNDQYPSGRSPPKDKHLFSLVK
EMRKAFBESVEKDIPLLSTGAGIIDVIKSGYKIPELSQSLDIOWMTYDLHDPK
DGYTGNSPLKSPYDVGKSDIANVDSIISYWKDHGAASEKLIYGFPRAYGHTFILSDP
SKTGIGATITSTGPPGKTDSEGLLAYIYEVCITFLNEGATEVNDAPQEVPIAYQGNWV
GYDNVRSPLKIAQWLKONNLGSAVWVPLMDMDFSGSFCHQRFHPLTSTLKGDLNLSA
SCKGPN"
sig_peptide 13..75
mat_peptide 76..1206
polyA_signal 1487..1492
polyA_site 1507
BASE COUNT 434 a 343 c 335 g 413 t
ORIGIN
Alignment Scores:
Pred. No.: 2,98e-95 Length: 1525
Score: 1387.00 Matches: 249
Percent Similarity: 80.8% Conservative: 55
Best Local Similarity: 66.22% Mismatches: 72
Query Match: 55.64% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-9 (1-452) x M94584 (1-1525)
Qy 1 TyrAsnLeuIleCysTrpPheThrAsnTrpAlaGlnTrpArgProGlyLeuGlySerPhe 20
Db 76 TACCAGCTGATGTGTACTATACCACTGGGTGGGTAAGGACAGGCCCAATAGAGGGAGTTTC 135
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleThrAlaPheAlaGlyMet 40
Db 136 AAACCTGGTAATATTAGCCCTGCTGTACTACCTACCTGATCTATGCTTGTCTGGAATG 195
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTrpLysAlaPheAsn 60
Db 196 CAGATATATGAGTACACTTACACATGACGACAGACTTGGCTGACTATGAGCATTTGAT 255
Qy 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 256 GGTCTGAAAGACAGACACACTGAGCTAAATAAACTCTCTCTGGCCATTGGAGGATGGAATTT 315
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 316 GGACCTGCCCCGTTCAGTGCCATGGTCTCTACTCTCTCAGAACCGTCAGATATTCATT 375
Qy 101 SerValIleLysPheLeuArgGlnTrpGlyPheAspGlyLeuAspLeuAspTrpGluTrp 120
Db 376 TCAGTTATCAGATTCCTTCGTCATATATACTTTGATGGCCTCAACCTGAGTGGCAGTAC 435
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 436 CTGGGTCTCGAGGAAGCCCTCTTAAGGACAAACATCTCTTCACTGTTCTGTGTGAAGAA 495
Qy 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 496 ATGCGTAAAGCTTTTGAAGGAAGAATCTCTGGAGAAAGACATTCACAGGCTGCTACT 555
Qy 161 AlaAlaValAlaGlyIleSerAsnIleGlnAlaGlyTrpGluIleProGluLeuSer 180
Db 556 TCCACAGGAGCAGGAGNATCATTTGACGTATCAAGTCTGGGTGACAGATCCCTGAATGCT 615
Qy 181 LysTyrLeuAspPheIleHisValMetThrThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 181 LysTyrLeuAspPheIleHisValMetThrThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
```



616 CAGTCTCTTGAATATATTCAGGTCATGACATATGATCTCCATGATCTTAAGGATGGCTAC 675  
201 ThrGlyGluAsnSerProLeuTyrIlystYrProThrGluThrGlySerAsnAlaTyrLeu 220  
676 ACTGGAGAAATAGTCCCTCTATATATCTCATGATGATGGAAGAGTCTGATCTC 735  
221 AsnValAspTyrValMetAsnTyrTrpIlyAsnAsnGlyAlaProAlaGluLysLeu 240  
736 AATGTGGATTCAATCATTTCTTACTTGAAGGACCATGAGGACGCTTCTGAGAAGCTCAT 795  
241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260  
796 GTGGGATTTCCAGCATATGCGCATACCTTTATCTCGAGTGACCTTCTTAAGACTGGAAT 855  
261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280  
856 GTGGCCCTACAAATAGTACTGCGCCACAGAAAGTACAGATGAAATCAGGACTCTCTG 915  
281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300  
916 GCTTACTATGAGGTTGTACATTTCTTGAATGAAGGAGCCACTGAGGTCGTGGATGCCCC 975  
301 GlnGluValProTyrAlaTyrIlyAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320  
976 CAGGAAGTACCTTATGCTTATCAGGATAATGAGTGGTGGTTGATGACATGTCAGGAGC 1035  
321 PheSerValIlyAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340  
1036 TTCAGTTGAAGGCTCAGTGGCTCAAGGACAAATTTAGGAGGTGCCGTGTGGCC 1095  
341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360  
1096 CTGGACATGATGATCTTCACTGCTTCTTCTGTCACGAGAGACATTTCCCTCTGACATCT 1155  
361 ThrLeuAsnIlyAlaLeuGlyLysThrGluGlyCysThrAlaPro 376  
1156 ACTTTAAGGAGATCTCAATATACACAGTGAAGTTGCAAGGGCCCT 1203

RESULT 12  
LOCUS D87757 1506 bp mRNA linear ROD 12-JAN-2000  
DEFINITION Mus musculus mRNA for ECF-L precursor, complete cds.  
ACCESSION D87757  
VERSION D87757.2 GI:6015435  
KEYWORDS ECF-L precursor.  
SOURCE Mus musculus bone marrow cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Ohashi, M., Arita, H. and Hayai, N.  
1 (bases 1 to 1506)  
Identification of a novel eosinophil chemotactic cytokine (ECF-L)  
as a chitinase family protein  
J. Biol. Chem. 275 (2), 1279-1286 (2000)  
2002902  
2 (bases 1 to 1506)  
Ohashi, M.  
Direct Submission  
Submitted (06-SEP-1996) Makoto Ohashi, The University of  
Tokushima, Faculty of Integrated Arts and Sciences; 1-1  
Minami-Johananjima, Tokushima, Tokushima 770, Japan  
(E-mail: ohashi@ias.tokushima-u.ac.jp, Tel: 0886-56-7261,  
Fax: 0886-56-7298)  
On Oct 7, 1999 this sequence version replaced gi:1545818.  
Sequence updated (05-Oct-1999).  
Location/Qualifiers  
1..1506  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue type="bone marrow"  
1..1506  
/gene="mECF-L"  
1..1197

/gene="mECF-L"  
/codon\_start=1  
/product="ECF-L precursor"  
/protein\_id="BAA13458.2"  
/db\_xref="GI:6015435"  
/translation="MAKULITGLAILNLVOLGSSYQLMCMYTSWAKORPIEGSPKPG  
NIDPLCLTHLYAFAGMNNIYTHQDLRDYEALNGDKNTKLTLLAIGWKKFG  
PASFSAMVSTPQNRQIFQISVIRFLQYVDFGLNDQYVPSRSGSPDKHLFSLVK  
EMRKAPESEVKEKDI PRLLTSTGAGI IDVIKSGYKI PELQSLDYI QVMTYDLHDPK  
DGYTGENSEPYSVDICKSADLNVDSIISYKOHGAASEKILVGFPAVGHFTFILSDP  
SKTGIGAPTISTGPCKTDESGLLAYYEVECTFLNEGATEVWDARQEVPAAYQNEWV  
GYDNVRSFKAKLMDKNNLGGVWVPLMDDDFSGSFQHRHFFPLTSLKGLNLHSA  
SCRGPY"  
BASE COUNT 423 a 338 c 333 g 412 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,17e-94 Length: 1506  
Score: 1379.00 Matches: 248  
Percent Similarity: 80.59% Conservative: 55  
Best Local Similarity: 65.96% Mismatches: 73  
Query Match: 55.31% Indels: 0  
DB: 10 Gaps: 0  
US-10-004-219B-9 (1-452) x D87757 (1-1506)  
QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20  
DB 64 TACCAGCTGATGTGCTACTATACCAGTTGGGCTAAGGACAGGCCAATAGAGGGAGTTTC 123  
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
DB 124 AAACCTGGTAATATGACCCCTGCTGTACTACCTGATCTATGCTTTCTGGAATG 183  
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrIlyAlaPheAsn 60  
DB 184 CAGAATAATGATGATCACTTACACATGACGACGAGACTTTCGCTGACTATGAAGCTTGAAT 243  
QY 61 AspLeuLysAsnAsnArgAsnSerLysLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80  
DB 244 GGTCTGAAGACAGACACACTGAGCTAAACCTCTCTGGCCATTTGGAGGATGGAAGTTT 303  
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100  
DB 304 GGACCTGCTCGTTCAGTGCCATGCTCTACTCTCTCAGAACCGTCAGATATTCATTTCAG 363  
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120  
DB 364 TCAGTTATCAGATTCCTTCGTCATATAAATCTTTGATGGCCTCAACCTGGACTGGCAGTAC 423  
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140  
DB 424 CTGGGCTCTGAGGAGGCCCTCTAAGGACAAACATCTCTTCAGTGTCTGTTGGTGAAGAA 483  
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160  
DB 484 ATGCGCTAAGCTTTTGGAGGAGAAATCTGTGGAGAGACATTCCAAGGCTGCTACTACT 543  
QY 161 AlaAlaValAlaGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
DB 544 TCCACAGGAGGAGGAATCATTCAGCTTAATCAAGTCTGGGTACCAAGATCCCTGAATCTCT 603  
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
DB 604 CAGTCTCTTGACTATATTCAGGTCATGATCATATGATCTCCATGATCTTAAGGATGGCTAC 663  
QY 201 ThrGlyGluAsnSerProLeuTyrIlystYrProThrGluThrGlySerAsnAlaTyrLeu 220  
DB 664 ACTGGAGAAATAGTCCCTCTATATAATCTCATGATGACATTTGGAAGAGTCTGATCTC 723  
QY 221 AsnValAspTyrValMetAsnTyrTrpIlyAsnAsnGlyAlaProAlaGluLysLeuIle 240  
DB 724 AATGTGGATTCAATCATTTCTTACTTGAAGGACCATGAGGAGCTTCTGAGAAGCTCAT 783

```
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
Db 784 GTGGGATTTCCAGCATATGGGCATACCTTTATCTCCTGAGTGACCTTCTAAGACTGGAAAT 843
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
Db 844 GTGCCCCCTACAAATAGTACTGGCCCCCAGGAGGAGTACACAGATGAATCAGGACTCCTG 903
QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 904 GCTTACTATGAGTTTGTACATCTTCTGAATGAAGGAGGACCACTGAGGTCTGGGATGCCCC 963
QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
Db 964 CAGGAAGTACCCTATCCCTATCAGGTAATGAGTGGTGGTTGTTATGACATGTGAGAGC 1023
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyAlaMetIleTrpAla 340
Db 1024 TTCAAGTTGAAGGCTCAGTGGCTCAAGGACAAACAATTTAGGAGGTGCGGTCTGGCCC 1083
QY 341 IleAspLeuAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
Db 1084 CTGGACATGAGTACTCAGTGGTCTTCTGTCTCACCAGAGACATTTCCCTCTGACATCT 1143
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaPro 376
Db 1144 ACTTTAAAGGAGATCTCAATATACACAGTGCAGTGCAGGCGCCT 1191

RESULT 13
LOCUS AY049765
DEFINITION Mus musculus secreted protein precursor Ym2 mRNA, complete cds.
ACCESSION AY049765
VERSION AY049765.2 GI:22123906
KEYWORDS house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1209)
Webb,D.C., McKenzie,A.N. and Foster,P.S.
Expression of the Ym2 lectin-binding protein is dependent on
interleukin (IL)-4 and IL-13 signal transduction: identification of
a novel allergy-associated protein
J. Biol. Chem. 276 (45), 41969-41976 (2001)
21551268
11553626
REFERENCE 2 (bases 1 to 1209)
Webb,D.C. and Foster,P.S.
Direct Submission
Submitted (02-AUG-2001) Biochemistry and Molecular Biology, John
Curtin School of Medical Research, Australian National University,
Off Mills Rd, Acton, Canberra, ACT 2601, Australia
3 (bases 1 to 1209)
Webb,D.C. and Foster,P.S.
Direct Submission
Submitted (06-AUG-2002) Biochemistry and Molecular Biology, John
Curtin School of Medical Research, Australian National University,
Off Mills Rd, Acton, Canberra, ACT 2601, Australia
Sequence update by submitter
On Aug 6, 2002 this sequence version replaced gi:15705156.
1..1209
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/note="isolated from the lung lavage from allergic mice"
1..1209
/note="expression regulated by IL-4 and IL-13 signaling
via the IL-4R alpha subunit"
/codon_start=1
/product="secreted protein precursor Ym2"
```

```
/protein id="AAL03953.2"
/db_xref="GI:22123907"
/translation="MAKILVLTGLAILNLVOLGSSYOLMCVYTWAKDRPTEGSPKFG
NIDCLCHLIYAFAGMKNNEITYLSEQLRDYELAGLKDRLTELKTLAAGWKKFG
PAPSSMVSTPQNRQTFIKSVIRFLRQDGLNDWOYPGSRGSPKDKHLFSLVQ
EMRAFEESTLNHIPLLSTGAGFIDVIKSGYKIPELSQSLDIQVMTYDHDPK
NGYTGNSPLYSYDIDKSLADNVDSITTYKDHGASERLIYGFPAHYHTFLSDP
SKNGIDPTVSGAGPPKYTNEQGLAYFEICTFLNEGATEIFDATQEPYAYLGNWV
CYDNVRSFKLKAQWLKONNLGAVVWPLDMDDFSSGFCHQGRFFPLTTTLKRDNLNVHSA
SCKASYRGL"
sig_peptide 1..63
BASE COUNT 326 a 273 c 282 g 328 t
ORIGIN
Alignment Scores: 2,57e-94 Length: 1209
Pred. No.: 1373.00 Matches: 245
Score: 81.60% Conservative: 61
Best Local Similarity: 65.33% Mismatches: 69
Query Match: 55.07% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-9 (1-452) x AY049765 (1-1209)
QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 64 TACCAGCTGATGCTGCTACTATACAGCTGGGTAAGGACAGGCCAACAGAGGGAGTTTC 123
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 124 AAACCTGTAATATATGACCCCTGCTGTGTACTCACCTGATCTATGCGCTTTGCTGGGATG 183
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 184 AAGAATAATGAGATCACCTTACTTAAGTGAGCAAGACCTTGGCGTACTATGACGATTAAT 243
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 244 GGTCTGAAAGACAGGACACACAGCTGAGCTAAACCTCTCTGGCCATTGGAGGATGGAAGTTT 303
QY 81 GlyThrAlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 304 GGACCTGCGCCGCTTCCAGTTCCTCTCTCTCTCAGAACCCGTCAGACATTCATTAAAG 363
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 364 TCAGTTATCAGATTCCTTCGTCATATTAACCTTTGATGCCCTCAACCTGCAGCTGGCGATAC 423
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 424 CCTGGGTCTCGAGGAGGCCCTCTTAAGGACAAACATCTCTTCAGTGTCTGCTGGCAGAA 483
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 484 ATGCGTAAAGCTTTTGGAGGAAGAATCCACTTTGAACCAACATTCACAGGCTGCTACTACT 543
QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 544 TCCACAGAGCTGGATTCATTGACGTAATCAGTCTGGGTACAGATCCCTGAATGCTGCT 603
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 604 CAGTCTCTCGACTATATTCAGGTCAATGATGATGATGATGATGATGATGATGATGATGATG 663
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 664 ACTGGAGAAAATAGTCCCTCTATAAATCTCATATGATGATGATGATGATGATGATGATG 723
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
Db 724 AATGGGATTCATTAATTTACTTACTTGGAGGAGGACCATGAGGAGGAGGAGGAGGAGGAG 783
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
```

```

|||||
784 GTGGGATTTCAGCATATGGTCTATACCTTTATCTCGAGTGACCCCTTCTAAGAATGGAATA 843
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGluAlaGlyPheTrp 280
Db 844 GTGACCCCTACTGTTAGTGTGCGACACACAGAAAGTACACAAATGAACAAGGATCTCTG 903
QY 281 AlaTyrTyrGluLeuCythrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 904 GCTTACTTTGAGATTGTGACATTTCTGAATGAAGAGAGCCACTGAGATCTTTGATGCCACC 963
QY 301 GlnGluValProTyrAlaTyrIysAlaGlnGluTrpLeuGlyTyrAspAsnIleYssSer 320
Db 964 CAGGAAGTACCTATGCTGCTATCTGGGTATGATGAGTGGTGGTTGTTATGACAAATGTCAGGAGC 1023
QY 321 PheSerValIysAlaGlnTrpLeuIysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
Db 1024 TTCAGTTGAAGGCTCAGTGGCTCAGGACAAATTTAGGAGGTGGCTGTCTGGGCC 1083
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
Db 1084 CTGGACATGGATGACTTCAGTGGTCTTCTCTCACCAGGAGCGTTTCCTCTGACAAC 1143
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAla 375
Db 1144 ACTTTAAAGAGAGATCTGTAATGTACACAGTGCAGAGTTGCAAGGCC 1188

```

## RESULT 14

```

BJA345054
LOCUS BJA345054 1562 bp mRNA linear VRT 14-JUN-2002
DEFINITION Bufo japonicus mRNA for toad pancreatic chitinase (tpCase gene).
ACCESSION AJ345054
VERSION AJ345054.1 GI:15963340
KEYWORDS toad pancreatic chitinase; tpCase gene.
SOURCE Japanese toad.
ORGANISM Bufo japonicus

```

## REFERENCE

```

1 Oshima, H.
AUTHORS Oshima, H.
JOURNAL Thesis (2001) Department of Department of Biology, Waseda
University, Tokyo, Japan
2 Oshima, H., Miyazaki, R., Ohe, Y., Hayashi, H., Kawamura, K. and
AUTHORS Kikuyama, S.
TITLE Isolation and sequence of a novel amphibian pancreatic chitinase
JOURNAL Comp. Biochem. Physiol. 132 (2), 381-388 (2002)
3 Oshima, H.
AUTHORS Oshima, H.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2001) Oshima H., Department of Biology, Waseda
University, Nishiwaseda 1-6-1, Shinjuku-ku, Tokyo, 169-8050, JAPAN
FEATURES
Location/Qualifiers
1..1562
/organism="Bufo japonicus"
/db_xref="taxon:8387"
/tissue_type="pancreas"
/country="Japan"
1..1562
/gene="tpCase"
2..1468
/gene="tpCase"
/EC_number="3.2.1.14"
/function="chitin fragmentation"
/codon_start=1
/evidence="experimental"
/product="toad pancreatic chitinase"
/protein_id="CAC87888.1"
/db_xref="GI:15963341"
/db_xref="SPTREMBL:Q90W34"
/translation="MLLWAGLFLLHLVQLGSGTYKLVVCYFTNWSQYRPDQKYPGNID
POLCTHLVYAFATMNEKIAPIYEWNDVLYKQFNDLKQKNKLVTLAIGWNFTQK

```

## gene

## CDS

```

FTDMVASSGNRSIFIKSVIAYLRQNNFDGIDLDFFEYSGSRGSPBEDKQRFVLIQEML
DAFNEEARSSGLPRLIITAASAGKGTIDAGYEIAKIQLLDFISVMTYDFHGWDTQ
SGHNSPLCKSGSTDYGLQYFNIHAMNWKNGGAPAEKLLGLLFTYGRFTFRPNPNC
DVGIIPVSGAGSAGPYTREAGWAYVEICTWLSGTQKWI PQRPVPAACKSNWVCFDN
QESYEKVRPLKESGFGGMVWADLDDEGFRFCNGRYPLINHLSKSLLEGSTVNCPC
EICGSIPTPTSATTTTTTTAKDCCTTPPEPVTTPPPVPIVDVDFNFCVEKTDGLH
VNPLNTNKFYICANGRTYSMKCADGLVFQASNCNCNP"
2..55
/gene="tpCase"
/evidence=experimental
56..1465
/gene="tpCase"
/product="toad pancreatic chitinase"
/EC_number="3.2.1.14"
/evidence=experimental
BASE COUNT 470 a 329 c 345 g 418 t
ORIGIN
Alignment Scores:
Pred. No.: 4,87e-94 Length: 1562
Score: 1371.00 Matches: 259
Percent Similarity: 65.42% Conservative: 55
Best Local Similarity: 53.96% Mismatches: 128
Query Match: 54.99% Indels: 38
DB: 5 Gaps: 6
US-10-004-219B-9 (1-452) x BJA345054 (1-1562)
QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 56 TACAAACTGGTGTGTACTTTACAAACTGGTCACAGTATCGCCCTGACGAGGAAGTAT 115
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 116 GTGCCAGGGAATATCGACCCGAGCTCTGCACACATCTCGTATATGCTTCGCCACCATG 175
QY 41 GlnAsnAsnGluIleThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 176 AATGAGCATAGATGCGCCATATGATGAATGAATGATGATGCTTATACAAACAGTTTAAAC 235
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuAlaIleGlyLysTrpAsnPhe 80
Db 236 GACCTAAAGCAGAAAATAAATAACCTGGTGACATTTATGGCAATCGGTGGTGAACATT 295
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 296 GGAACACAGAGATTTCACAGATATGTTCTCTCTGGAACCGGTCAATTTTCATTAA 355
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 356 TCTGTGATTGCATATCTACGACAAATAATTTTGTATGAATAGACCTTGATTTTGAATAT 415
QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 416 CCAGGATCTCGAGGAAGTCTCCAGAGGATAAGCAGCGATTACTGTACTGATACAGAA 475
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 476 ATGCTTGATGCAATTAATGAAGAAGCAAGATCTCTCGATGATGATAGATAGCAAGAA 535
QY 161 AlaAlaValAlaGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 536 GCAGCAGTGTGACGAGGAAAAGGCCACCATCGATGCTGATGATGATAGATAGCAAGAA 595
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 596 CAATTGCTGGATTTTATCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 655
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 656 TCTGACACACAGCCCTCTGTGCAAGAGTTCCACTGACTACGAGAGATTGCAATACTTT 715
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240

```

Db	716	AAcATTcATTtTCCcATgATTtTGGAAaAaCAAGGGGcCCcCTGcAGAAaACTcCTT	775
Qy	241	ValGlyPheProGluTyrGlyHisThrPheLeuArgAsnProSerAspAsn	258
Db	776	CTTGGATtTCCcACATgATGcCGcAAcATTc-----AGAAcCCAaTCTTAcATGtGT	829
Qy	259	GlyIleGlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGly	278
Db	830	GATGTGGGATcCCcGTGTCTGGAGcAGATcAGTGGcCCATAcACcAGAGGcGTGC	889
Qy	279	PheTrpAlaTyrTyrGluLeuCysThrPheLeuArgSerGlyAlaThrGluValTrpAsp	298
Db	890	TTCTGGcCTAcTATcAGATcTCTGcATTGGcCT---TCAGGAGcACAGTGAaATGgATT	946
Qy	299	AlaSerGlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIle	318
Db	947	CCAGACcAGAGATGcCATATGCTTCcAAAGATTAAGTGGGTTCGATTGcACACCAa	1006
Qy	319	LysSerPheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIle	338
Db	1007	GAAAGTATcGAATcGAAGTcAGTCTTAAAGAAAGTGGGTTCGTGGcCATGgTT	1066
Qy	339	TrpAlaIleAspLeuAspPheThrGlySerPheCysAspGlnGlyLysPheProLeu	358
Db	1067	TGGcCCATtGATTGGATGATTTCGAGGcGAGATTtTGCaATcAAGGAAGATcCGTtG	1126
Qy	359	ThrSerThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCys-----	373
Db	1127	ATCAcACcCTGAAGTCTTGTGGAGGTTCaACTGTAaCTGGcCTCGAGAAATATGT	1186
Qy	374	-----	Thr 374
Db	1187	GGTGGcATTCTATAcCAACTcCTCAAGTGCcAACcACCACcACcACcAGcC	1246
Qy	375	AlaProAsp-----ValProSerGluProValThrThrProProGlySerGly	392
Db	1247	AAcCTGcACTGcACTAcCAGAGcCTcCAGTcCAGTcCCTcCTcCAGcT-----	1291
Qy	393	GlyGlySerSerGlyGlySerGlyGlySerGlyPheCysAlaAspLysAlaAspGly	412
Db	1292	-----GTcCTcCTcCTcATGATGATGTTGACcCCcAACTcCTGtGTAGAAaACTGATGc	1345
Qy	413	LeuTyrProValAlaAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyr	432
Db	1346	CTTcATGTAaTCCcACTAAATAcAAaCAAAATtTTAcATTtTGcAAATGcAGTAcCTAT	1405
Qy	433	GlnGlnHisCysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrpPro	452
Db	1406	TCAATGAaATGTGcAGATGcCTTAGTATTTCAGGcCAAGTTCcAACTGTTGAcTGcCA	1465
RESULT 15			
LOCUS	AY065557	1209 bp	mRNA linear ROD 22-MAY-2002
DEFINITION	Mus musculus putative secretory protein precursor (Ym2) mRNA, complete cds.		
ACCESSION	AY065557		
VERSION	AY065557.1	GI:18086513	
KEYWORDS			
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Chang, N.-C.A.		
TITLE	Tissue-specific Expression of Ym2 Protein		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1209)		
AUTHORS	Chang, N.-C.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-DEC-2001) Institute of Microbiology & Immunology, National Yang-Ming University, #155 Sec. 2 Li-Nong St., Beitou, Taipei 112, Taiwan		

FEATURES		Location/Qualifiers	
source		1..1209	
		/organism="Mus musculus"	
		/strain="BALB/c"	
		/db_xref="taxon:10090"	
		/clone="pGEMT-ez-ym2.s"	
		/tissue type="stomach"	
gene		1..1209	
		/gene="Ym2"	
CDS		1..1209	
		/gene="Ym2"	
		/codon_start=1	
		/product="putative secretory protein precursor"	
		/protein_id="AAL57751.1"	
		/db_xref="GI:18086514"	
		/translation="MAKILIVTGLAILNLVOLGSSYQLMVCYVTSWAKDRPTGSGFKPG NIDPCLTHLIYAFAGMKNNKITLSEQDLRDYDALNGLKDRNTKTLKLLAIGWKFG PAPSSVMSTPQNRQTFIKSVIRFLQRYDFGLNDLMOYPGRSGPPKDLFLSVLQ EMKAFEBESTLNHPRLLTSTGAGTIDVYSGYKIPELSQSLDYIQVMTYDLHDPK NGYTGNSPLKSPYDYGKADLVDSIIITWKDHGAASEKLVGFPAFGHTFILSDP SKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDQTOEVPYAVLGNWV GYDVRSFKLKAQWLKONNLGAVVPLMDMDFSGSFCHQGRFPLITTLKRDNLNVHSA SCKASVYRGEL"	
BASE COUNT		326 a	274 c 282 g 327 t
ORIGIN		US-10-004-219B-9 (1-452) x AY065557 (1-1209)	
Alignment Scores:		Pred. No.: 6,11e-94 Length: 1209	
		Score: 1368.00 Matches: 244	
		Percent Similarity: 81.33% Conservative: 61	
		Best Local Similarity: 65.07% Mismatches: 70	
		Query Match: 54.87% Indels: 0	
		DB: 10 Gaps: 0	
Qy	1	TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe	20
Db	64	TACCAGCTGATGTGTACTATACAGCTGGCTAAGGACGAGCCAAAGAGGGAGTTTC	123
Qy	21	LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet	40
Db	124	AAACCTGTATATATGACCCCTGCTGTACTACCTGATCTATGCTTGTGGGATG	183
Qy	41	GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn	60
Db	184	AAGATATATGATGACCTTACTTAAGTCAGCAGCACTTGGCTGACTATGAAGCATTAA	243
Qy	61	AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe	80
Db	244	GGTCTGAAAGACAGCAACACTGAGCTAAAAACTCTCTCTGGCCATTGGAGGATGGA	303
Qy	81	GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr	100
Db	304	GGACCTGCCCGCTTCAGTTCCTACTCTCTCAGAACCCGTCAGACATTCAATTAAG	363
Qy	101	SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGlyTyr	120
Db	364	TCAGTTATCAGATTCCTTCGTCATATACTTGTAGTGCCTCAACCTGAGCTGAGTAC	423
Qy	121	ProGlySerArgGlySerProGlnAspLysHisLeupheThrValLeuValLysGlu	140
Db	424	CCTGGCCCTCAGGAGAGCCCTCTTAAGGACAAACATCTCTTCACTGTTCTGGTGAG	483
Qy	141	MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr	160
Db	484	ATGCGTAAAGCTTTTGAGGAAGAATCCACTTTTGAACCACTTCCAAGGCTGCTACT	543
Qy	161	AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer	180
Db	544	TCCACAGGAGCTGGATTTCATTCAGCTTAATCAAGCTCGGTACAGATCCCTGACTGT	603
Qy	181	LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr	200



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:02:46 ; Search time 248.065 Seconds  
(without alignments)  
4103.372 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLICYFTNWAQYRPLGLSF.....QQHCQAGLVFDTSCNCCNWP 452

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2195239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xl  
-Q/cgn2\_1/USFTO\_spool/US10004219/runat\_24062003\_160227\_8151/app.query.fasta\_1.1877  
-DB=N Geneseq 101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CGN 1 1 626 @runat\_24062003\_160227\_8151 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	2089.5	83.8	1368	22	AAH42013	Disease treatment
2	2089.5	83.8	1678	22	AAH42025	Disease treatment
3	1778.5	71.3	1369	24	ABNS9993	Novel human coding
4	1387	55.6	1469	22	AAH42023	Disease treatment
5	1301.5	52.2	1636	19	RAV10435	Human chitinase c1
6	1301.5	52.2	1636	20	AAD21847	MO-218 clone of hu
7	1301.5	52.2	1636	22	AAD03759	Human chitinase cd
8	1301.5	52.2	1636	24	ABL57380	Human chitinase cd
9	1301.5	52.2	1643	18	AAT50833	Human chitinase cd
10	1295.5	52.0	1656	19	AAV10436	Human chitinase c1
11	1295.5	52.0	1656	20	AAZ21848	MO-13B Clone Of hu
12	1295.5	52.0	1656	22	AAD03760	Human chitinase cd
13	1295.5	52.0	1656	24	ABL57381	Human chitinase cd
14	1294.5	51.9	1768	18	AAT89181	Human chitotriosid
15	1286	51.6	1713	18	AAT50834	Human chitinase cd
16	1279.5	51.3	1637	18	AAT89180	Human chitotriosid
17	1051	42.2	2504	16	AAQ90443	Murine oviduct spe
18	1033.5	41.5	2366	16	AAQ90444	Hamster oviduct sp
19	1010	40.5	1994	16	AAQ90442	Bovine oviduct spe
20	1008	40.4	2346	23	ABL03597	Drosophila melanog
21	971.5	39.0	1681	16	AAQ85245	YKL-40 gene. Homo
22	971.5	39.0	1681	21	AAZ94901	Human cancer marke
23	963.5	38.6	1594	19	AAV21688	DNA encoding a hum
24	956.5	38.4	1152	18	AAT99452	Bovine whey protei
25	955.5	38.3	1432	19	AAV21689	DNA encoding a hum
26	955.5	38.3	1594	19	AAV21687	Drosophila melanog
27	954.5	38.3	17368	23	ABL03596	Drosophila melanog
28	950.5	38.1	1433	19	AAV13925	Human cartilage gp
29	950.5	38.1	1496	18	AAV9127	Human cartilage g1
30	950.5	38.1	1526	19	AAV13926	cDNA sequence #402
31	902.5	36.2	1391	24	AA63615	Drosophila melanog
32	879.5	35.3	1433	23	ABL15099	Drosophila melanog
33	864	34.7	3552	23	ABL15098	Drosophila melanog
34	863.5	34.6	2452	18	AAT62557	Manduca sexta larv
35	854.5	34.3	1474	22	AAH23078	Osteoarthritis tis
36	851	34.1	13497	23	ABL03557	Drosophila melanog
37	830.5	33.3	2510	23	ABL06628	Drosophila melanog
38	782	31.4	2089	23	ABL29841	Drosophila melanog
39	775	31.1	966	9	AAH81756	Gene encoding poly
40	773.5	31.0	638	24	ABK11713	DNA encoding novel
41	772.5	31.0	746	22	AAK86635	Human digestive sy
42	761.5	30.5	16489	23	ABL03556	Drosophila melanog
43	756	30.3	1302	23	ABL19745	Drosophila melanog
44	756	30.3	3302	23	ABL19744	Drosophila melanog
45	729.5	29.3	1608	21	AAZ38579	D. farinae mite al

ALIGNMENTS

RESULT 1  
AAH42013  
ID AAH42013 standard; DNA; 1368 BP.  
XX AC AAH42013;  
XX DT 24-AUG-2001 (first entry)  
XX DE Disease treatment related oligonucleotide SEQ ID NO: 3.  
XX KW Disease treatment; infection; chronic occlusive pulmonary disease;  
XX KW bronchial asthma; ds.  
XX OS Homo sapiens.  
XX PN WO200136633-A1.  
XX PD 25-MAY-2001.







XX  
PR 11-SEP-2000; 2000US-0659671.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX WPI; 2002-292408/33.  
DR P-PSDB; ABB97580.  
DR  
XX  
PT An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis -  
XX  
PS Claim 1; SEQ ID NO 404; 509pp; English.  
XX  
CC The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a coding sequence of the  
CC invention.

XX SQ Sequence 1369 BP; 327 A; 393 C; 341 G; 308 T; 0 other; .

Alignment Scores:

Pred. No.:	2.88e-150	Length:	1369
Percent Identity:	1778.50	Matches:	315
Best Local Similarity:	89.14%	Conservative:	38
Query Match:	71.34%	Mismatches:	40
Indels:	24	Gaps:	1

US-10-004-219B-9 (1-452) x ABN59993 (1-1369)

QY	59	PheAsnAspLeuLysAAsArgAnSerLysLysLeuLysThrLeuLeuAlalleGlyGlyTrr	78
Db	67	TATACTGAATTTCAGACCTCGGAACAGCAGCTGAAAACTCTCTGGCCCATTTGGAGGCTGG	126
QY	79	AnsPheGlyThrAlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPhe	98
Db	127	AAC TTCGGGAGCTGCCCTTTCTAGCCCATGTGTTTCTCTCTTGAGAACCAGCACTTTC	186
QY	99	IleThrSerValIleLysPheLeuArgGlnTyrglyPheAspGlyLeuAspLeuAspTrp	118
Db	187	ATCACCTCAGTCATCAAATTCCTGC GCCAGATGAGTTTGACGGGCTGGACTTTGACTGG	246
QY	119	GluTyrrProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuVal	138
Db	247	GAGTACCTCGCTCTGTGGGAGCCCTCTCTCAGGACAAGCATCTCTTCACTGTCTCTGGTG	306
QY	139	LysGluMetArgGluAlaPheGluCluAlalleGluSerAsnArgProArgLeuMet	158
Db	307	CAGGAAAATGCGTAAGCTTTTTGAGCAGAGAGCCCAAGCAGATCAACAAGCCAGGCTGATG	366
QY	159	ValThrAlaAlaValAlaGlyLysSerAsnIleGlnAlaGlyTyrrGluIleProGlu	178
Db	367	GTCACTGTGCAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCA	426
QY	179	LeuSerLysTyrrLeuAspPheIleHisValMetThrTyrrAspLeuHisGlySerTrpGlu	198
Db	427	CTGTCCACAGTACTCGGACTACATCATGTCATGACCTACGACCTCCATGGCTCTCTGGAG	486
QY	199	GlyTyrrThrGlyGluAsnSerProLeuTyrrLysTyrrProThrGluThrGlySerAsnAla	218
Db	487	GGCTACACTGGAGAGAACAGGCCCCCTCTACAAAATACCCGACTACACCGGACCAACGCC	546
QY	219	TyrrLeuAsnValAspTyrrValMetAsnTyrrTrpLysAsnGlnValProAlaGluLys	238

XX WPI; 2001-397791/42.  
 XX New proteins, peptides and DNA for treatment of bronchial asthma,  
 PT chronic occlusive lung disease and infectious disease  
 XX Disclosure; Page 106; 114pp; Japanese.  
 XX The present invention provides the sequence of a protein which can be  
 CC used in the treatment and prevention of infectious diseases. Inhibitors  
 CC of the protein can be used to treat bronchial asthma and chronic  
 CC occlusive pulmonary disease. The present sequence is an oligonucleotide  
 CC described in the exemplification of the invention.  
 XX Sequence 1469 BP; 406 A; 331 C; 332 G; 400 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 4,866-115 Length: 1469  
 Score: 1387.00 Matches: 249  
 Percent Similarity: 80.85% Conservative: 55  
 Best Local Similarity: 66.22% Mismatches: 72  
 Query Match: 55.64% Indels: 0  
 DB: 22 Gaps: 0

US-10-004-219B-9 (1-452) x AAH42023 (1-1469)

QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20  
 DB 72 TACCAGCTGATGTCTACTATACCTAGTGGCTAGGACAGGCCAATAGAGGGAGTTTC 131

QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
 DB 132 AAACCTGGTAATATTGACCCCTGCTGTACTACCTGTATCTGCTGGAATG 191

QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrIleAlaPheAsn 60  
 DB 192 CAGAATAATGAGATCCTTACACATGATGACGAGCAAGCTTGGCTATGAGCATTTGAAT 251

QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPhe 80  
 DB 252 GGTCTGAAAGACAGACAGACTGAGCTTAACCTCTCTGGCCATTTGGAGATGGAGTTT 311

QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100  
 DB 312 GGACCTGCCCTTCAGTCCCATGTCTCTACTCTCTCAGAACCGTCAGATATTCATTCAG 371

QY 101 SerValIleLysPheIleuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120  
 DB 372 TCAGTTATCAGATTCCTTCGTCATATATACCTTTGATGGCTCAACCTGGACTGGCAGTAC 431

QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140  
 DB 432 CTGGGTCTCGAGGAAGCCCTCTAGGACAAACATCTCTCAGTGTCTGGTGAAGGAA 491

QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160  
 DB 492 ATGCGTAAAGCTTTTGAGGAGAACTGTGGAGAAAGACATTCCAAGGCTGCTACTACT 551

QY 161 AlaAlaValAlaGlyClyIleSerHenIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
 DB 552 TCCACAGGAGCAGGAATCATACGCTAATCAAGTCTGGGTACAGATCCCTGAACCTGCT 611

QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
 DB 612 CAGTCTCTTGACTATATTCAGGTGATGATATGATATGATATGATATGATATGATATGAT 671

QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220  
 DB 672 ACTGGAGAAAATAGTCCCTCTATAAATCTCCATATGATGATTTGGAAGAGTCTGATCTC 731

QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240  
 DB 732 AATGTGGATTCAATTTCTTACTGGAAGGCCATGAGCTTCTGAGAAGCTCATTT 791

QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260  
 DB 792 GTGGATTTCCAGCATATGGGCATACCTTTATCTCTGAGTGACCTTCTTAAGACTGGAATT 851

QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280  
 DB 852 GGTGCCCTTACAATTTAGTACTGGCCACCAGGAAAGTACACAGATGAATCAGGACTCTCTG 911

QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300  
 DB 912 GCTTACTATAGGTTTGTACATTTCTGAATGAAGAGGCCACTGAGGTCTGGATGCCCCC 971

QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320  
 DB 972 CAGGAAGTACCTTATGCTATCAGGGTAATCAGTGGTGGTTGTTATGACAAATGTCCAGGAC 1031

QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340  
 DB 1032 TTCAGTTTGAAGGCTCAGTGGCTCAAGGACCAACAATTTAGGAGGTGCGGTGCTGGGCC 1091

QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360  
 DB 1092 CTGGACATGGATGACTTCAGTGGTTCCTTCTGTCTACCAGAGACATTCCTCTGACATCT 1151

QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaPro 376  
 DB 1152 ACTTTAAGGGAGATCTCAATATACACAGTGCAGTGCAGGGGCCCT 1199

RESULT 5  
 AAV10435  
 ID AAV10435 standard; cDNA; 1636 BP.  
 XX AAV10435;  
 AC AAV10435;  
 XX 15-JUN-1998 (first entry)  
 DT Human chitinase clone MO-218 cDNA.  
 DE Chitinase; human; fungal infection; immunogen; diagnosis; treatment;  
 KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;  
 KW Rheumatoid arthritis; overexpression; extracellular matrix; ss.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FT CDS 2..1402  
 FT /\*tag= a  
 FT sig\_peptide 2..64  
 FT /\*tag= b  
 FT mat\_peptide 65..1399  
 FT /\*tag= c  
 FT /product= chitinase  
 FT /note= "from clone MO-218"  
 XX WO9747752-A1.  
 PN 18-DEC-1997.  
 PD 16-JUN-1997; 97WO-US10460.  
 PF 14-JUN-1996; 96US-0663618.  
 PR (ICOS-) ICOS CORP.  
 PA Gray PW;  
 PI WPI; 1998-052316/05.  
 DR P-PSDB; AAH40259.  
 XX Nucleic acids encoding human chitinase - useful as antifungal  
 PT agents, especially in combination with other antifungals  
 XX

Claim 3; Page 38-40; 63pp; English.

This sequence encodes a novel human chitinase isolated from clone MO-218. Chitinases are useful for treating or preventing fungal infection and as immunogens for generating antibodies which are used to purify, detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The nucleic acid sequence of the chitinase is also useful as a probe to identify and isolate genomic DNA encoding chitinases or similar proteins, or cells expressing them or to generate transgenic ('knockout') rodents. It can also be used in hybridisation assays and to detect genetic alterations in the chitinase gene related to disease. Agents that inhibit this protein may be useful in treatment of Gaucher's disease and rheumatoid arthritis, where overexpression of the protein can damage the extracellular matrix. Chitinase also improves the activity of other antifungal agents and may allow a reduction in the dose of such agents, and thus of their side effects.

Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

#### Alignment Scores:

Pred. No.:	2,71e-107	Length:	1636
Score:	1301.50	Matches:	235
Percent Similarity:	67.11%	Conservative:	71
Best Local Similarity:	51.54%	Mismatches:	129
Query Match:	52.21%	Indels:	21
DB:	19	Gaps:	3

US-10-004-219B-9 (1-452) x AAV10435 (1-1636)

QY	3	LeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro	22
DB	71	CTGGTCTGTCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCTCTGCC	130
QY	23	AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn	42
DB	131	AAGGACTTGGACCCAGCCCTTGGACCCACCTATCTACGCTTCTGCTGGCATGACCAAC	190
QY	43	AsnGluileThrThrileGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu	62
DB	191	CACCAGCTGAGCACCACTGAGTGGATGACGAGACTCTTACCAGAGTTCAATGGCCTG	250
QY	63	LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyGlyTrpAsnPheGlyThr	82
DB	251	AAGAATGAATCCCAAGCTGAAGACCTGTAGCCATCGAGGCTGGAATTTCCGCACT	310
QY	83	AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheThrSerVal	102
DB	311	CAGAACTTCAAGATATGTGTAGCCAGGCCCAACACCGTCAGACCTTTGTCAACTCGGCC	370
QY	103	IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly	122
DB	371	ATCAGGTTCTGGCAATACAGCTTTGACGGCTTGACCTTGACCTGGAGTACCAGGA	430
QY	123	SerArgLysSerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg	142
DB	431	AGCCAGGGAGCCCTCCGCTAGCAAGAGGCGCTTCAACACCTGTGTACAGACTTGGCC	490
QY	143	GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla	162
DB	491	ATGCTCTCCAGCAGGAAGCCAGACTCAGGAAGGAACCGCTTCTCTGAGTGGAGCG	550
QY	163	ValAlaGlyGlyIleSerAsnileGlnAlaGlyTyrGluileProGluLeuSerLysTyr	182
DB	551	GTTCAGCTGGGCAGACCTATGTGGATGTGGATACAGGTGGACAAATCGCCAGAAC	610
QY	183	LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly	202
DB	611	CTGGATTTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTACGGGA	670
QY	203	GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal	222
DB	671	CATAACAGCCCCCTCTACAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCTCAACGFG	730

QY	223	AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly	242
DB	731	GATGCTGTGTGCAACAGTGGCTGCAGAGGGGAGCCCTGCAGCAAGCTGATCTTGGC	790
QY	243	PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla	262
DB	791	ATGCTTACTTACGAGCGCTCTTCACTGGCTCTCATCAGACACCCAGAGTGGGGGCC	850
QY	263	ProThrSerGlyAspGlyProAlaGlyAlaTyrArgGlnAlaGlyPheTrpAlaTyr	282
DB	851	CCAGCCACAGGCTGGGACCTCCAGGCCCTTCACCAAGGAGGAGGATGCTGGCTTAC	910
QY	283	TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu	302
DB	911	TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG	964
QY	303	ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer	322
DB	965	GTGCCCTCATCTTCCGGGCAACACAGTGGGTGGGCTTTGATGTGGAGAGCTTCAAA	1024
QY	323	VallLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp	342
DB	1025	ACCAAGCTCAGCTATCTAGCAGAGGAGGACTGGGGGGGCGCATGCTGGGCACTGGAC	1084
QY	343	LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu	362
DB	1085	TTAGATGACTTTGGCGGCTTCTCTGCAACACGAGGCGCATACCCCTCATCCAGACGCTA	1144
QY	363	AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal-----	378
DB	1145	CGGAGGAACTAGTCTTCCATATCTTGGCTTTCAGGACCCAGAGCTTGAAGTTCCAAA	1204
QY	379	-----ProSerGluProValThrThrProProGlySerGlySerGlyGlySer	395
DB	1205	CCAGCTCAGCCCTCGAACCT-----	1228
QY	396	SerGlyGlySerSerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro	415
DB	1229	CATGCCCCCAGCCCTGGACAAGACACAGCTTCTGCCAGGCAAGCTGATGGGCTCTATCC	1288
QY	416	ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis	435
DB	1289	AATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCGGGGGCGGCTGTTCAGCAAGAC	1348
QY	436	CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp	451
DB	1349	TGCCCGACAGGCGCTGTGTTCAGCAACTCTCTGCAATGCTGCACCTGG	1396

#### RESULT 6

AAZ21847  
ID AAZ21847 standard; DNA; 1636 BP.

XX AAZ21847;

AC AAZ21847;

XX 10-DEC-1999 (first entry)

DE MO-218 clone of human Chitinase, with noncoding 5'/3' regions.

XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;

XX organ transplant; parasite; chitin-binding; allele; vector;

XX truncated protein; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 2..1402

XX /tag= a

XX /product= Human\_Chitinase

XX sig\_peptide 2..65

XX /tag= b

XX /product= Signal\_peptide

XX mat\_peptide 66..1402

XX /tag= c

/product= Mature\_protein

FT XX W09946390-A1.  
 PN XX 16-SEP-1999.  
 PD XX 12-MAR-1999; 99WO-US05343.  
 PF XX 12-MAR-1998; 98US-0039198.  
 PR XX (ICOS-) ICOS CORP.  
 PA XX Gray PW, Tjoelker LW;  
 PI XX WPI; 1999-551417/46.  
 DR XX P-PSDB; AY42425.  
 DR XX  
 PT Novel chitin-binding fragments of human chitinase used to treat fungal  
 PT infections in animals  
 PS Example 1; Page 55-57; 83pp; English.  
 XX This is the nucleotide sequence of an allelic form of the human  
 CC chitinase enzyme, which is capable of degrading Chitin (a linear  
 CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).  
 CC Chitinase fragments can be used to screen for proteins or other  
 CC molecules that specifically bind to the chitin-binding domain of human  
 CC chitinase or that modulate its activity. These compounds are useful for  
 CC immunization, as well as for purifying chitinase, as well as for  
 CC detection and quantification of chitinase. Polynucleotide fragments of  
 CC the invention are useful as a source of probes and primers, and to  
 CC express the proteins recombinantly. The chitinase fragments, when  
 CC conjugated to antifungal compounds, are used to treat animals,  
 CC especially humans, infected with chitin-containing parasites such as  
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
 CC sporotrichosis, and dermatophycoses.  
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
 CC protein for treating infections, especially fungal infections, is  
 CC problematic. In view of the increasing incidents of life-threatening  
 CC fungal infection in e.g. immunocompromised individuals, there exists a  
 CC need for identifying new compounds for treating fungal infection. The  
 CC chitin-binding fragments of the present invention provide this need.

XX SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,71e-107 Length: 1636  
 Score: 1301.50 Matches: 235  
 Percent Similarity: 67.11% Conservative: 71  
 Best Local Similarity: 51.54% Mismatches: 129  
 Query Match: 52.21% Indels: 21  
 DB: 20 Gaps: 3

US-10-004-219B-9 (1-452) x AA221847 (1-1636)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
 DB 71 CTGGTCTGCTACTTACCAACTGGGCCCAAGTACAGACAGGGGGAGCGCTCGCTTCCTGCC 130  
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42  
 DB 131 AAGGACTTGGAGCCCGCCAGCTTTGGACCCCACTCACTACGCTTCCTGGCATGACCAAC 190  
 QY 43 AsnGluileThrTrileGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
 DB 191 CACCACTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGGAGTTCAATGSCCTG 250  
 QY 63 LysAsnArgAsnSerLysLeuileThrLeuAlaileGlyTyrTrpAsnPheGlyThr 82  
 DB 251 AAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGGAGGCTGGAATTTTCGGCACT 310

QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
 DB 311 CAGAAGTTTCACAGATATGGTAGCCCAACCGCTCAGACCTTTGTCAACTCGGCC 370  
 QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
 DB 371 ATCAGGTTTCTGCGCAATACAGCTTTGACGCGCTTGACCTTGAGTGGGAGTACCAGGA 430  
 QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
 DB 431 AGCCAGGGAGCCCTGCGTAGACAGAGGCGCTTCACACCTGGTACAGGACTTGGCC 490  
 QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162  
 DB 491 AATGCTTTCAGCAGGAAGCCAGACCTCAGGAAGGAACGCTTCTTCTGAGTGCAGCG 550  
 QY 163 ValAlaGlyGlyLysSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
 DB 551 GTTCCAGCTGGGCGACCTATGTGATCTGATACGAGGTGGACAAATCGCCAGAAC 610  
 QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGly 202  
 DB 611 CTGATTTTGTCAACCTTATGCGCTTACGACTTCCATGCTCTTGGGAGAGTCAAGGA 670  
 QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
 DB 671 CATAACAGCCCTCTACAGAGCAAGAGAGTGTGCGAGCAGCCAGCCTCAACGTG 730  
 QY 223 AspTyrValMetAsnTyrTrpLysAsnGlyAlaProAlaGluLysLeuileValGly 242  
 DB 731 GATGCTGTGTCAACAGTGTGCGAAGGGGAGCCCTGCGCAGCAAGCTATCTCTGGC 790  
 QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
 DB 791 ATGCTTACTACGAGCGCTCTCTACCTGGCTCTCATCAGCAGCAGGAGTGGGGCC 850  
 QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
 DB 851 CCAGCCACAGGCTGTGGCACTCCAGGCCCTTCCACCAAGGAAGGAGGATGCTGGCTAC 910  
 QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
 DB 911 TATGAAGTCTGCTCTCTGG-----AAGGGGGGCCCAACACAGAGAATCCAGGATCAGAAG 964  
 QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
 DB 965 GTGCCCTACATCTTCCGGGCAACCCAGTGGGTGGCTTTGATGATGGAGAGTTCAAA 1024  
 QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaileAsp 342  
 DB 1025 ACCAAGTTCAGCTATCTGACAGCAAGAGGAGTGGGGGGGCCCATGCTCTGGGCACTGGAC 1084  
 QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
 DB 1085 TTAGATGACTTTGCGCGCTTCTCTGCAACACAGGGGCCGATACCCCTCATCCAGACGCTA 1144  
 QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378  
 DB 1145 CGGCAGGAAGTCTGCTTCCATCTTCCATCTTCCAGCAGCCCGAGCTTGAAGTTCACAAA 1204  
 QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlyGlySer 395  
 DB 1205 CCAGTTCAGCCCTCTGAACCT-----GAG 1228  
 QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415  
 DB 1229 CATGGCCCCAGCCCTGGCAAGACACAGCTTCTGCCAGGGCAAGAGCTGATGGGCTATCC 1288  
 QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435  
 DB 1289 AATCTCGGAAACGGTCTCAGCTTCTACAGCTGTGCGGGGGCGGCTGTTCAGCAAGAC 1348  
 QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451

Db 1349 TGCCGACAGGCGCTGGTGTTCAGCACTCTCGAAATGCTGCACCTGG 1396

## RESULT 7

AD03759

ID AAD03759 standard; cDNA; 1636 BP.

AC AAD03759;

XX 19-JUN-2001 (first entry)

XX Human chitinase cDNA from clone pMO-218.

DE Human; antifungal; chitinase; immunoglobulin; Ig; therapy;

XX fungal infection; candidiasis; aspergillosis; coccidioidomycosis;

KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;

KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;

KW clone pMO-218; ss.

XX Homo sapiens.

OS Key

XX Location/Qualifiers

PH 2..1402

FT /\*tag= a

FT /product= "Human chitinase from clone pMO-218"

FT sig\_peptide 2..64

FT /\*tag= b

FT mat\_peptide 65..1399

FT /\*tag= c

FT /product= "Human mature chitinase from clone pMO-218"

FT W0200123430-A2.

PN 05-APR-2001.

XX 28-SEP-2000; 2000MO-US26960.

XX 30-SEP-1999; 99US-0409918.

XX (ICOS-) ICOS CORP.

XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;

XX WPI; 2001-266141/27.

XX P-PSDB; AA200432.

XX Novel chitinase immunoglobulin fusion product, useful for treating

PT fungal infections and reducing the amount of a non-chitinase antifungal

PT agent needed for the treatment

XX Claim 2; Page 29-31; 39pp; English.

XX The present invention relates to a chitinase immunoglobulin (Ig) fusion

CC product, comprising a human chitinase fused to at least a portion of an

CC immunoglobulin chain. The fusion product is useful for treating fungal

CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,

CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,

CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.

CC The fusion protein is useful for reducing the amount of non-chitinase

CC antifungal agent needed to exert an antifungal activity. The fusion

CC protein is also useful for preparing a medicament for the prophylactic

CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin

CC fusion product has unexpectedly improved serum half-life and formulation

CC properties. The present sequence is human chitinase cDNA from clone

CC pMO-218. Chitinase enzyme degrades chitin which is a homopolymer of

CC beta-(1,4)-linked N-acetylglucosamine residues.

XX Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

SQ

Alignment Scores:

Pred. No.: 2,71e-107 Length: 1636

Score: 1301.50 Matches: 235

Percent Similarity: 67.11% Conservative: 71

Best Local Similarity: 51.54% Mismatches: 129  
Query Match: 52.21% Indels: 21  
DB: 22 Gaps: 3

US-10-004-219B-9 (1-452) x AAD03759 (1-1636)

QY 3 LeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
DB 71 CTGGTCTGTACTTCCACCACTGGGCCCCAGTACAGACAGGGGGAGGCTCGCTTCTGCCC 130  
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAen 42  
DB 131 AAGGACTTGGACCCCGCCCTTTGGACCCACCTCATCTAGCCCTTCGCTGGGATGACCAAC 190  
QY 43 AsnGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
DB 191 CACCACTGAGCACCACCTGAGTGGATGACGAGACTCTCTACCAAGAGTTCAATGGCCTG 250  
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPheGlyThr 82  
DB 251 AAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGGAGGCTGGAATTTCCGCACT 310  
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
DB 311 CAGAAATTCACAGATATGTTAGCCACGCGCAACACCTCTAGACCTTTGTCAACTCGGCC 370  
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
DB 371 ATCAGTTTCTGGCAATATACAGCTTTGACGCGCTTGACTTGGAGTACCCAGGA 430  
QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
DB 431 AGCCAGGGGAGCCCTGCGGTAGACAAAGAGGCGCTTCAACACCTGGTACAGACTTGGCC 490  
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162  
DB 491 AATGCTTTCAGCAGGAGGCCAGACCTCAGAGGAAGAACCGCTTCTTCTAGTGCAGCG 550  
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
DB 551 GTTCCAGCTGGGCAGACCTATGTGATGCTGGATAGAGGTGACAAATGCCCCAGAAC 610  
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
DB 611 CTGGATTTGTCACTTATGCGCTTACGACTTCCATGCTCTTGGGAGAAGTCCAGGGA 670  
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
DB 671 CATAACACCCCTCTACAGAGGCAAGAGAGTGTGTCAGCAGCGCTCAACCTG 730  
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242  
DB 731 GATGCTGTGTGCAACAGTGGCTGCAGAGGGGCCCTTCCAGCAAGCTGATCTTGGC 790  
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
DB 791 ATGCTACTACTCGGACGCTCTCTACACTGGCTCTCATCGACACCAAGAGTGGGGGCC 850  
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
DB 851 CCAGCCACAGAGTCTGGGCATCTCCAGGCCCTTCCAAAGGAGGAGGATGTCGCTCCATAC 910  
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
DB 911 TATGAGTCTGCTCTCTGG-----AAGGGGGCCCAACAGAGATCCAGGATCAGAAG 964  
QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
DB 965 GTGCCCTACATCTTCCGGGCAACACCACTGGTGGTGGCTTTGATGATGTGGAGAGCTTCAA 1024  
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342  
DB 1025 ACCAAGGTTCAGCTATCTGAAGCAGAGGAGTGGCGGGGGCCATGCTCTGGCACTTGAC 1084

QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
 |||||  
 DB 1085 TTAGATGACTTTGGCGGTTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGCTA 1144  
 |||||  
 QY 363 AnLysAlaLeuGlyLysSerThrGluGlyCysThrAlaProAspVal----- 378  
 |||||  
 DB 1145 CGGCAAGTCTGAGTCTTCATCTTCCATCTGCTTCAGGCACCCAGAGCTTGAAGTTCCAAAA 1204  
 |||||  
 QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395  
 |||||  
 DB 1205 CCAGTGCACCTCTGAACCT-----GAG 1228  
 |||||  
 QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyPro 415  
 |||||  
 DB 1229 CATGGCCCCAGCCCTGGACACACAGCTTCTGCGAGGCGAAAGCTGATGGGCTCTATCCC 1288  
 |||||  
 QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyTrpGlnGlnHis 435  
 |||||  
 DB 1289 ANTCTCGGGAAGGTCAGCTTCTACAGCTGTGCAGCGGGCGGCTGTTCAGCAAGC 1348  
 |||||  
 QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTyr 451  
 |||||  
 DB 1349 TGCCCGACAGGCTGGTGTTCAGCACTCTGCAAAATGCTGCACCTGG 1396  
 |||||

## RESULT 8

ABL57380  
 ID ABL57380 standard; cDNA; 1636 BP.

AC ABL57380;

DT 12-AUG-2002 (first entry)

DE Human chitinase cDNA clone MO-218.

XX Chitinase; enzyme; human; fungicide; antifungal; infection;  
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;  
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
 KW Pneumocystis; gene; ss.

XX Homo sapiens.

FF Key Location/Qualifiers  
 FT CDS 2..1402  
 FT /\*tag= a  
 FT /product= "Chitinase"  
 FT sig\_peptide 2..64  
 FT /\*tag= b  
 FT mat\_peptide 65..1399  
 FT /\*tag= c

PN US6372212-B1.

XX 16-APR-2002.

XX 16-JUN-1997; 97US-0877599.

XX 14-JUN-1996; 96US-0663618.

XX (ICOS-) ICOS CORP.

XX Gray PW;

XX WPI; 2002-442449/47.

XX P-PSDB; ABB76291.

XX Co-administering chitinase to improve the effectiveness of fungicidal  
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -  
 XX Example 1; Column 23-28; 26pp; English.

CC The present sequence is that of cDNA clone MO-218 (ATCC 98077)  
 CC encoding human chitinase (see ABB76291). The clone was isolated  
 CC from a cDNA library prepared from peripheral blood monocyte-derived  
 CC macrophages following sequence analysis. Also isolated was clone  
 CC MO-13B (see ABL57381), which contains a single nucleotide difference  
 CC in the coding region, changing the encoded amino acid at position 81  
 CC of the mature protein from glycine to serine. Northern blots showed  
 CC highest chitinase gene expression in lung and ovary tissues.  
 CC Expression in lung is consistent with a protective role against  
 CC pathogenic organisms that contain chitin. The invention provides  
 CC human chitinase polynucleotides and polypeptides, and materials and  
 CC methods for the recombinant production of human chitinase products,  
 CC which are expected to be useful as products for treating fungal  
 CC infections or for the development of such products. Human  
 CC chitinase has a synergistic effect on the actions of other  
 CC fungicides. It can be administered to improve the antifungal  
 CC activity of a non-chitinase antifungal agent, especially  
 CC amphotericin B or itraconazole, in the treatment of a fungal  
 CC infection such as candidiasis, aspergillosis, coccidioidomycosis,  
 CC blastomycosis, paracoccidioidomycosis, histoplasmosis,  
 CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,  
 CC dermatophytoses and Pneumocystis infections (all claimed). In  
 CC particular, the fungal infection involves Candida, Aspergillus  
 CC and/or Cryptococcus spp., whose growth is not effectively  
 CC inhibited by contact with human chitinase alone.

XX Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

## Alignment Scores:

Pred. No.: 2,71e-107 Length: 1636  
 Score: 1301.50 Matches: 235  
 Percent Similarity: 67.11% Conservative: 71  
 Best Local Similarity: 51.54% Mismatches: 129  
 Query Match: 52.21% Indels: 21  
 DB: 24 Gaps: 3

US-10-004-219b-9 (1-452) x ABL57380 (1-1636)

QY 3 LeuileCysTyThrPheThrAsnTrpAlaGlnTyArgProGlyLeuGlySerPheLysPro 22  
 |||||  
 DB 71 CTGGTCTGCTACTTCCCAACTGGGCCCCAGTACACAGGGGGAGGCTCGTTCCTGCC 130  
 |||||  
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyTrpAlaPheAlaGlyMetGlnAsn 42  
 |||||  
 DB 131 AAGGACTTGGACCCAGCCTTTGCACCACCTCATCTAGCCCTTCGTCGCATGACCAAC 190  
 |||||  
 QY 43 AsnGluileThrThrIleGluTrpAsnAspValThrLeuTyLysAlaPheAsnAspLeu 62  
 |||||  
 DB 191 CACCAGCTGAGCACCACCTGAGTGGTAATGACGAGACTCTCTACCGAGGTTCATATGGCCTG 250  
 |||||  
 QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTyTrpAsnPheGlyThr 82  
 |||||  
 DB 251 AGAAGATGATCCCAAGCTGAAGACCTGTGTAGCANTCGGAGGCTGGANTTCGGCAT 310  
 |||||  
 QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
 |||||  
 DB 311 CAGAAGTTTCACAGATATGGTAGCCACGCCCAACAACCGTCAGACCTTTGTCACATCGGCC 370  
 |||||  
 QY 103 IleLysPheLeuArgGlnTyTrpGlyPheAspGlyLeuAspLeuAspTrpGluTyTrpGly 122  
 |||||  
 DB 371 ATCAGGTTTCTCGCAATAATACAGCTTTGACGCCCTTGACCTTGACTGGGAGTACCCAGGA 430  
 |||||  
 QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
 |||||  
 DB 431 AGCCAGGGGAGGCCCTGCCGTAGCAAGGAGGCGCTTTCACAACCCCTGGTACAGGACTTGGCC 490  
 |||||  
 QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162  
 |||||  
 DB 491 AATGCTTCCAGCAGAGAGCCAGACCTCAGGAGGAGGAGCGCTTCTTCTGAGTGCAGCG 550  
 |||||  
 QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyTrpGluileProGluLeuSerLysTyr 182  
 |||||  
 DB 551 GTTCCAGCTGGGCAGACCTATGTGGATGCTGGATACGAGGTGACAAATATCGCCAGAAC 610  
 |||||

QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
 DB 611 CTGGATTGTCAACCTTATGGCTACGACTCCATGGCTCTTGGGAGAGGTTCACGGGA 670  
 QY 203 GluAsnSerProLeuTyrIleTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
 DB 671 CATACAGAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGCAGCAGCTTCAACGTG 730  
 QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuValGly 242  
 DB 731 GATGCTGTGTCAACAGTGGCTGCAGAGGGAGCCCTGCGCAGCAGCTGATCTTGGC 790  
 QY 243 PheProGluTyrGlyHisThrPheLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
 DB 791 ATGCTTACCTACGAGCGCTCTTCACTGGCTCTCATCAGACAGAGTGGGGCC 850  
 QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
 DB 851 CCAGCCACAGGGTCTGGCACTCCAGGCCCTTCAACCAAGAGAGGAGTGTCTGGCCTAC 910  
 QY 283 TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
 DB 911 TATGAAGTCTGCTCTGG-----AAGGGGGGCCCAACCAAGAGAGTGTCTGGCCTAC 964  
 QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
 DB 965 GTGCCCTACATCTCCGGGACACCACTGGTGGTGGCTTGTATGATGTGGAGCTTCAAA 1024  
 QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnGlyGlyAlaMetIleTrpAlaIleAsp 342  
 DB 1025 ACCAAGTCTAGTATCTGAAGCAGAGAGGACTGGGGGGGCGCATGCTCTGGCCTGGAC 1084  
 QY 343 LeuAspAsnPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
 DB 1085 TTAGATGACTTGGCCGGCTTCTCTGCAACCAAGGGCGGATACCCCTCATCAGAGCTA 1144  
 QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal-----378  
 DB 1145 CGGCAGGACTGAGTCTTCCACTTTCCTTCCAGGACCCAGAGCTTGAAGTTCACAAA 1204  
 QY 379 -----ProSerGluProValThrThrProProGlySerGlyGlyGlySer 395  
 DB 1205 CCAGGTCAGCCCTCTGAAGCT-----GAG 1228  
 QY 396 SerGlyGlySerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415  
 DB 1229 CATGGCCCGCAGCCCTGGCAAGACAGCTTCTGCGAGGCAAGCTGATGGCTCTATCCC 1288  
 QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435  
 DB 1289 AATCCTCGGGAGCGTCCAGCTCTTACAGCTGTGCAGCGGGGGCGGTGTTCAGCAAGC 1348  
 QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451  
 DB 1349 TGCCCGCAGAGCGCTGTGTTCAGCAACTCTGCAATGTGTCACCTGG 1396  
 RESULT 9  
 AAT50833  
 ID AAT50833 standard; cDNA; 1643 BP.  
 XX AC AAT50833;  
 XX DT 24-MAR-1997 (first entry)  
 XX DE Human chitinase cDNA clone chi.50.  
 XX KW Chitinase; chitotriosidase; chitin; infectious disease;  
 KW gene therapy; vaccine; lysosomal lipidoses; Gaucher's disease;  
 KW leishmaniasis; sarcoidosis; x-linked adrenoleukodystrophy;  
 KW multiple sclerosis; drug delivery; cosmetics; food; ss.  
 OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT CDS 13..1413  
 FT /\*tag= a  
 FT sig\_peptide 13..75  
 FT /\*tag= b  
 FT mat\_peptide 76..1410  
 FT /\*tag= c  
 XX WO9640940-A2.  
 XX 19-DEC-1996.  
 XX 06-JUN-1996; 96WO-NL00225.  
 XX 07-JUN-1995; 95US-0486839.  
 XX (UNAM ) UNIV AMSTERDAM.  
 XX Aerts JMF; PI  
 XX WPI: 1997-118698/11.  
 XX P-PSDB; AAW08584.  
 XX New human chitinase - used to treat or prevent infection by  
 PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,  
 PT foods, implants etc.  
 XX Claim 2; Page 39-40; 58pp; English.  
 XX A cDNA clone (AAT50833), designated chi.50, codes for a 50 kDa human  
 CC chitinase (AAW08584) that is stable to many proteases, active at pH 3-  
 CC 8 and up to 50 deg, and stable in the circulation. Clones chi.50  
 CC and chi.39 (see also AAT50834) were isolated from a human macrophage  
 CC cDNA library using as probe a partial clone obtd. using primers  
 CC (see also AAT50835-36) based on a chitotriosidase purified from a  
 CC type 1 Gaucher disease patient. The 2 cDNA clones are the result  
 CC of alternative splicing of RNA. Chitinase nucleic acid can be used  
 CC for large-scale prodn. of recombinant human chitinases, or can be  
 CC incorporated into a gene therapy vector to treat or prevent  
 CC infection by chitin-contg. pathogens.  
 XX Sequence 1643 BP; 364 A; 490 C; 442 G; 347 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2,73e-107 Length: 1643  
 Score: 1301.50 Matches: 235  
 Percent Similarity: 67.11% Conservative: 71  
 Best Local Similarity: 51.54% Mismatches: 129  
 Query Match: 52.21% Indels: 21  
 DB: 18 Gaps: 3  
 US-10-004-219B-9 (1-452) x AAT50833 (1-1643)  
 QY 3 LeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
 DB 82 CTGGTCTGCTACTTCCACCACTGGGCCCCAGTACAGAGGGGAGGCTCGCTTCGTGCC 141  
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42  
 DB 142 AAGGACTTGGAGCCCGAGCCCTTTCACCCACCTCATCTACGCTTCGCTGGCATGCCAAC 201  
 QY 43 AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
 DB 202 CACGAGCTGAGCACCACCTGAGTGGATGACAGAGACTCTCTACGAGGACTTCATGGCCTG 261  
 QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyTyrTrpAsnGlyThr 82  
 DB 262 AAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGGAGGCTGGAAATTCGGCACT 321  
 QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
 DB 322 CAGAAGTTTCCAGATATGTTAGTCCAGCGGCCCAACACCGCTGAGACCTTTGTCAACTCGGCC 381



Qy	103	IleIysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly	122
Db	382	ATCAGGTTTCTGGCGAAATACAGCTTTGACGGCTTGACCTTGACTGGAGTAGTACCAGGA	441
Qy	123	SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg	142
Db	442	AGCCAGGGAGCCCTGCCGTAGACAGAGGCGCTTACAAACCTTGGTACAGGAGTTGGCC	501
Qy	143	GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla	162
Db	502	AATGCCCTTCAGCAGGAAGCCAGACCTCAGGGAAGGAAAGCCCTTCTTCGTGAGTCAGCG	561
Qy	163	ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr	182
Db	562	GTTTCCAGCTGGCGAGACCTATGTGGATGTGGATACGAGGTGACAAATATCGCCCGAAG	621
Qy	183	LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly	202
Db	622	CTGGATTTTGTCAACCTTATGGCTACGACTTCATGGCTCTTGGGGAAGGTTCACGGGA	681
Qy	203	GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal	222
Db	682	CATAACAGCCCCCTCTCAAGAGACGAAGAGAGAGTGGTGCAGCAGCCCTCAACGTG	741
Qy	223	AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProHlaGluLysLeuIleValGly	242
Db	742	GATGCTCTGTGCACAGTGGCTGCAGAGGGGACCCCTGCCAGCAGCTGATCTTTGGC	801
Qy	243	PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla	262
Db	802	ATGCTACTTACGAGACGCTCTTACACTGGCTCTCTCATACACCAGCAGTGGGGGCC	861
Qy	263	ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr	282
Db	862	CCAGCCACAGGGTCTGGCACTCCAGGCCCTTCCACCAAGAGAGGAGTGTGGCTTAC	921
Qy	283	TyrGluIleCysThrPheLeuArgSerGlyAlaThrCluValTrpAspAlaSerGlnGlu	302
Db	922	TATGAAGTGTCTCTCTGG-----AAGGGGGCCACCAAAACAGAGAAATCCAGGATCAGAAG	975
Qy	303	ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer	322
Db	976	GTGGCTTACATCTCCGGGACACCACTGGTGGCTTTCATGATGTGGAGAGCTTCAA	1035
Qy	323	ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp	342
Db	1036	ACCAAGGTTCAGTATCTGAAGACAGAGGGACTGGCGGGGCCATGTGTCTGGGCATGGAC	1095
Qy	343	LeuAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu	362
Db	1096	TTAGTACATTTGGCGGCTTCTCTCGCAACAGGGCGGATACCCCTTCATCCAGACGCTA	1155
Qy	363	AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal-----	378
Db	1156	CGGCAGGAACTGAGTCTTCCATCTTCCCTTCAGGCACCCAGAGCTTGAAGTTCACAAA	1215
Qy	379	-----ProSerGluProValThrThrProProGlySerGlyGlyGlySer	395
Db	1216	CCAGGTTCAGCCCTCTGAACCT-----GAG	1239
Qy	396	SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro	415
Db	1240	CATGGCCCCAGCCCTGCACAAAGACACGTTCTGCCAGGGCAAGCTGATGGGCTCTATCCC	1299
Qy	416	ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis	435
Db	1300	AATCCTCGGGAACGTCCTCAGCTTCTACAGCTGTGCAGCGGGCGGCTGTTCACGAAAGC	1359
Qy	436	CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp	451
Db	1360	TGCCCGACAGGCTGTGGTTTCAGCAACTCTCTGCAAAATGCTGCACCTGG	1407

```

RESULT 10
AAV10436
ID AAV10436 standard; CDNA; 1656 BP.
AC
XX AAV10436;
XX
XX
XX
XX
XX 15-JUN-1998 (first entry)
XX
XX
XX Human chitinase clone MO-13B cDNA.
XX
XX Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
XX KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
XX KW rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 27..1427
XX /*tag= a
XX
XX sig_peptide 27..89
XX /*tag= b
XX
XX mat_peptide 90..1424
XX /*tag= c
XX
XX /product= chitinase
XX /note= "from clone MO-13B"
XX
XX
XX WO9747752-A1.
XX
XX
XX 18-DEC-1997.
XX
XX
XX 16-JUN-1997; 97WO-US10460.
XX
XX
XX 14-JUN-1996; 96US-0663618.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW;
XX
XX
XX WPI: 1998-052316/05.
XX P-PSDB; AN40260.
XX
XX
XX Nucleic acids encoding human chitinase - useful as antifungal
XX PT agents, especially in combination with other antifungals
XX
XX
XX Claim 9; Page 42-44; 63pp; English.
XX
XX
XX This sequence encodes a novel human chitinase isolated from clone MO-13B.
XX CC Chitinases are useful for treating or preventing fungal infection and
XX CC as immunogens for generating antibodies which are used to purify, detect
XX CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
XX CC nucleic acid sequence of the chitinase is also useful as a probe to
XX CC identify and isolate genomic DNA encoding chitinases or similar proteins,
XX CC or cells expressing them or to generate transgenic ('knockout') rodents.
XX CC It can also be used in hybridisation assays and to detect genetic
XX CC alterations in the chitinase gene related to disease. Agents that inhibit
XX CC this protein may be useful in treatment of Gaucher's disease and
XX CC rheumatoid arthritis, where overexpression of the protein can damage
XX CC the extracellular matrix. Chitinase also improves the activity of other
XX CC antifungal agents and may allow a reduction in the dose of such agents,
XX CC and thus of their side effects.
XX
XX
XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

```

Qy	3	LeulleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro	22
Db	96	CTGGTCTGCTACTTTCAACAACCTGGGCCCCAGTACAGACAGGGGAGGCGTCTGCTTCTCTGCCC	155
Qy	23	AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn	42
Db	156	AAGACATTGACCCACCCAGCCTTTGACCCACCTCATCTACGCCCTTCGTGTCGATGACCAAC	215
Qy	43	AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu	62
Db	216	CACCAGCTGAGCACCACTGAGTGGATGACGAGACTCTCTACCAAGGAGTTCAATGGCCTG	275
Qy	63	LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr	82
Db	276	AAGAAGATGAATCCCAAGCTCAAGACCCTGTAGCCATCGAGGCTGGAAATTTTCAGCACT	335
Qy	83	AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal	102
Db	336	CAGAAGTTTCAGATATGGTAGCCACCGCCCAACAACCGTCAGACCTTTGTCAACTCGGCC	395
Qy	103	IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly	122
Db	396	ATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGACTGGGAGTACCCAGGA	455
Qy	123	SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg	142
Db	456	AGCCAGGGGAGCCCTGCGGTAGACAAGGAGCGCTTCACAACCCCTGGTACAGGACTTTGGCC	515
Qy	143	GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla	162
Db	516	AATGCCCTTCAGCAGGAAGCCAGACCTCAGGGAAGGAACGCCCTTCTTCGATGTGACGG	575
Qy	163	ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr	182
Db	576	GTTCACGCTGGGACAGACTATGTGGATGTGGATACGAGTGGACAAATAATCGCCAGAAC	635
Qy	183	LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly	202
Db	636	CTGGATTTTGTCAACCTTATGGCCTACGACTTCCTCCATGGCTTTGGGAGAAGGTCAACGGGA	695
Qy	203	GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal	222
Db	696	CATAACAGCCCCCTTCAAGAGAGCAAGAGAGAGTGTGCAGCAGCCACGCCCTCAACAGTG	755
Qy	223	AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGly	242
Db	756	GATCCTGCTGTGCACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTCTTGGC	815
Qy	243	PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla	262
Db	816	ATGCCCTACCTACGGACGCTCCTTCACACTGGCCTCCTCATCACACACAGAGTGGGGGCC	875
Qy	263	ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr	282
Db	876	CCAGCCACAGGGTCTGGCACTCCACAGGCCCTTCCACCAAGGAAGAGGAGTGTGGCCCTAC	935
Qy	283	TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu	302
Db	936	TATGAAGTGTGCTCCTCG-----AAGGGGGCCACCAACAGAGATCCAGATCCAGATCAGAG	989
Qy	303	ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer	322
Db	990	GTGCCCTACATCTTCGGGACAACACAGTGGGTGGGCTTTCATGATGTGGAGAGCTTCAAA	1049
Qy	323	ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp	342
Db	1050	ACCAAGGTACGACTATCTGAAGACAGAAGGAGTGGGGCGGGCCATGTGTCTGGCAGCTGGAC	1109
Qy	343	LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu	362
Db	1110	TTGATGACTTTGCGCGCTTCTCTCGAACACAGGGCCGATACCCCTTCATCCAGACGCTA	1169
Qy	363	AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal-----	378

Db	1170	CGCAGAACTGAGTCTTCATACTGCTTCAGGCACCCAGAGCTTGAGTTCACAAA	1229
Qy	379	-----ProSerGluProValThrThrProProGlySerGlySerGlyGlySer	395
Db	1230	CCAGGTCAGCCCTCTGAACCT	1253
			-----GAG
Qy	396	SerGlyGlySerSerGlyGlySerGlyPheCysAlaAaspGlyLeuTyPro	415
Db	1254	CATGGCCCCAGCCCTGGCAAGACAGCTTCTGCAGGGCAAGCTGATGGGCTCTATCCC	1313
Qy	416	ValAlaAaspArgAsnAlaPheTrpGlnCysIleAsnGlyLeuThrTyGlnGlnHis	435
Db	1314	AATCCTGGGAACGGTCCAGCTTCTACAGCTGTCAGCGGGCGGCTGTTCCAGCAAGC	1373
Qy	436	CysGlnAlaGlyLeuValPheAspThrSerCysAasnCysCysAasnTrp	451
Db	1374	TGCCCGACAGCCCTGGTGTTTCAGCAACTCTCTCAAAATGCTGCACCTGG	1421
RESULT 11			
ID	AAZ21848		
AC	AAZ21848 standard; DNA; 1656 BP.		
XX	AAZ21848;		
XX	10-DEC-1999 (first entry)		
DT	XX		
DE	MO-13B clone of human Chitinase, with noncoding 5'/3' regions.		
KW	chitin; fungal infection; immunocompromised; AIDS; chemotherapy;		
KW	organ transplant; parasite; chitin-binding; allele; vector;		
KW	truncated protein; ds.		
XX	XX		
XX	Homo sapiens.		
XX	XX		
Key	Location/Qualifiers		
FT	27..1427		
FT	/*tag= a		
FT	/product= Human_Chitinase		
FT	27..89		
FT	/*tag= b		
FT	/note= "Signal peptide"		
FT	90..1427		
FT	/*tag= C		
FT	/note= "Mature peptide"		
XX	XX		
PN	WO9946390-A1.		
XX	XX		
PD	16-SEP-1999.		
XX	XX		
PF	12-MAR-1999; 99WO-US05343.		
XX	XX		
PR	12-MAR-1998; 98US-0039198.		
XX	XX		
PA	(ICOS-) ICOS CORP.		
XX	XX		
PI	Gray, PW, Tjoelker LW;		
XX	XX		
DR	WPI; 1999-551417/46.		
DR	P-PSDB: AAY42426.		

Novel chitin-binding fragments of human chitinase used to treat fungal infections in animals -  
Example 1; Page 59-62; 83pp; English.  
This is the nucleotide sequence of an allelic form of the human chitinase enzyme, which is capable of degrading Chitin (a linear homo polymer of Beta-1,4-linked N-acetylglucosamine residues). Chitinase fragments can be used to screen for proteins or other molecules that specifically bind to the chitin-binding domain of human chitinase or that modulate its activity. These compounds are useful for immunization, as well as for purifying chitinase, as well as for

CC detection and quantification of chitinases. Polynucleotide fragments of  
 CC the invention are useful as a source of probes and primers, and to  
 CC express the proteins recombinantly. The chitinase fragments, when  
 CC conjugated to antifungal compounds, are used to treat animals,  
 CC especially humans, infected with chitin-containing parasites such as  
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
 CC sporotrichosis, and dermatophytoses.  
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
 CC protein for treating infections, especially fungal infections, is  
 CC problematic. In view of the increasing incidents of life-threatening  
 CC fungal infection in e.g. immunocompromised individuals, there exists a  
 CC need for identifying new compounds for treating fungal infection. The  
 CC chitin-binding fragments of the present invention provide this need.

XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

## Alignment Scores:

Pred. No.: 9.53e-107 Length: 1656  
 Score: 1295.50 Matches: 234  
 Percent Similarity: 66.89% Conservative: 71  
 Best Local Similarity: 51.32% Mismatches: 130  
 Query Match: 51.97% Indels: 21  
 DB: 20 Gaps: 3

US-10-004-219b-9 (1-452) x AAZ21848 (1-1656)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
 DB 96 CTGGTCTGCTACTTACCACAACTGGGCCCGGAGTACAGACAGGGGGGAGGCTCGCTTCCTGCC 155  
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42  
 DB 156 AAGGACTTGGACCCCGAGCTTTGGACCCACCTACTACCTTCTGCTGGCATGACCAAC 215  
 QY 43 AsnGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
 DB 216 CACCACTGAGCACCACTGAGTGGGAATGACGAGACTCTCTACCAAGAGTTCAATGGCCTG 275  
 QY 63 LysAsnArgAsnSerLysLeuLysThrLeuAlaIleGlyGlyTyrAsnPheGlyThr 82  
 DB 276 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGCTGGAAATTCAGCACT 335  
 QY 83 AlaProPheThrThrMetValSerThrGlnAsnArgGlnThrPheIleThrSerVal 102  
 DB 336 CAGAAGTTCACAGATATGTTAGCCAGCGGCCAACACACCTTCAGACCTTTGTCACTGGCC 395  
 QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
 DB 396 ATCAGGTTTCTGGCAAAATACAGCTTTGACGGCTTGACCTTGACCTGGGAGTACCCAGGA 455  
 QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
 DB 456 AGCCAGGGGAGCCCTGCCGTAGACAGGAGCGCTTCACAAACCTCGTACAGACTTGGCC 515  
 QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162  
 DB 516 AATGCTTCCACAGAGAACCCAGACCTCAGGAGGAGAACCGCTTCTCTGATGGCAGG 575  
 QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182  
 DB 576 GTTCCAGCTGGGCAGACCTATGTGGATGTGTGATGACGAGGTGGACAAATCGCCAGAAC 635  
 QY 103 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
 DB 636 CTGGATTTTGTCAACCTTATGGCCTACGACTTCCATGGCTCTTGGGAGAAAGTCCAGGGA 695  
 QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
 DB 696 CATACAGCCCCCTCTACAAAGAGGCGCAAGAGAGTGGTGGCAGCGCCCTCAACGTG 755  
 QY 223 AspTyrValMetAsnTyrTrpLysAsnGlyAlaProAlaGluLysLeuileValGly 242

DB 756 GATGCTGTGTGCAACAGTGGCTGCAGAAAGGGAGCCCTGCGCAGCAAGCTGATCTTGGC 815  
 QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
 DB 816 ATGCTTACTACGAGCGCTCTCTCACACTGCTCTCATCAGACACCCAGAGTGGGGGCC 875  
 QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
 DB 876 CCAGCCACAGGGTCTGGCACTCCAGGCCCTTCACCAAGGAGGAGGATGCTGGCCTAC 935  
 QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
 DB 936 TATGAAGTCTGCTCTGG-----AAGGGGGCCCAACCAACAGAGAATCAGGATCAGAAG 989  
 QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
 DB 990 GTGCCCTACATCTTCCGGGACACACGAGTGGGGCTTTGATGATGTGGAGAGTTCAAA 1049  
 QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342  
 DB 1050 ACCAAGGTCACTATCTGACGACAGAGGAGGACTGGGGCGGGCCATGCTTGGCACTGGAC 1109  
 QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
 DB 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACACGAGGCGCATACCCCTCATCCAGACGCTA 1169  
 QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378  
 DB 1170 CGGCAAGAACTAGTCTTCCATACCTTCCAGCACCCAGAGCTTGAAGTTCCAAA 1229  
 QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlyGlySer 395  
 DB 1230 CCAGGTCAAGCTCTGAACCT-----GAG 1253  
 QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415  
 DB 1254 CATGCCCCAGCCCTGGACACAGACACAGTCTTCCAGGCGCAAGCTGTATGCTTATCCC 1313  
 QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435  
 DB 1314 AATCTCGGGAACGGTCCAGCTTCTACAGCTGTGAGCGGGGGCGGCTGTTCAGCAAAAGC 1373  
 QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451  
 DB 1374 TGCCCGCAGGCGCTGTCTTCCAGCAACTCTCTGCAATGCTGCACCTGG 1421  
 RESULT 12  
 AAD03760  
 ID AAD03760 standard; cDNA; 1656 BP.  
 XX  
 AC AAD03760;  
 XX  
 DT 19-JUN-2001 (first entry)  
 XX  
 DE Human chitinase cDNA from clone pMO-13B.  
 KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;  
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;  
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;  
 KW clone pMO-13B; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 Key Location/Qualifiers  
 CDS 27..1427  
 FT /\*tag= a  
 FT /product= "Human chitinase from clone pMO-13B"  
 FT sig\_peptide 27..89  
 FT /\*tag= b  
 FT mat\_peptide 90..1424  
 FT /\*tag= c

/product= "Human mature chitinase from clone pMO-13B"

FT XX WO200123430-A2.  
 PN XX 05-APR-2001.  
 XX XX 28-SEP-2000; 2000WO-US26960.  
 PF XX 30-SEP-1999; 99US-0409918.  
 PR XX (ICOS-) ICOS CORP.  
 PA XX  
 PI Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;  
 DR WPI: 2001-266141/27.  
 DR P-PSDB; ABE00433.  
 XX  
 PT Novel chitinase immunoglobulin fusion product, useful for treating  
 PT fungal infections and reducing the amount of a non-chitinase antifungal  
 PT agent needed for the treatment -  
 XX  
 XX Claim 2; Page 34-36; 39pp; English.  
 PS  
 CC The present invention relates to a chitinase immunoglobulin (Ig) fusion  
 CC product, comprising a human chitinase fused to at least a portion of an  
 CC immunoglobulin chain. The fusion product is useful for treating fungal  
 CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,  
 CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
 CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.  
 CC The fusion protein is useful for reducing the amount of non-chitinase  
 CC antifungal agent needed to exert an antifungal activity. The fusion  
 CC protein is also useful for preparing a medicament for the prophylactic  
 CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin  
 CC fusion product has unexpectedly improved serum half-life and formulation  
 CC properties. The present sequence is human chitinase cDNA from clone  
 CC pMO-13B. Chitinase enzyme degrades chitin which is a homopolymer of  
 CC beta-(1,4)-linked N-acetylglucosamine residues.  
 XX  
 SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:  
 Pred. No.: 9,53e-107 Length: 1656  
 Score: 1295.50 Matches: 234  
 Percent Similarity: 66.89% Conservative: 71  
 Best Local Similarity: 51.32% Mismatches: 130  
 Query Match: 51.97% Indels: 21  
 DB: 22 Gaps: 3

US-10-004-219B-9 (1-452) x AAD03760 (1-1656)

QY 3 LeuileCysTyrPheThrAsnTrpAlaGlnTyrAtgProGlyLeuGlySerPheLysPro 22  
 DB 96 CTGGTCTGTACTTACCACTGGGCCAGTACAGACAGGGAGGCTCGCTTCCTGCC 155  
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42  
 DB 156 AAGGACTTGGACCCAGCCTTTGACCCACCTCATCTACGCTTCGCTGGCATGCCAAC 215  
 QY 43 AsnGluileThrThrIleGlnTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
 DB 216 CACCAGCTGAGCACCCTCAGTGGATGATGACGAGACTCTACAGGAGTTCAATGGCCCTG 275  
 QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyGlyTyrAsnPheGlyThr 82  
 DB 276 AGAAGATGAATCCCAAGCTGAGACCTTTGATGCCATCGGAGCTGGAAATTCAGCACT 335  
 QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheileThrSerVal 102  
 DB 336 CAGAAAGTTCACAGATATGTTAGCCAGCGGCCAACACCGTCAGACCTTTGTCAACTCGGCC 395  
 QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGlnTyrProGly 122  
 DB 396 ATCAGGTTTCTCGCAAAATACAGCTTTGACGGCTTGACCTTGACTGGGAGTACCAGGA 455

QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
 DB 456 AGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCACACCCCTGGTACAGGACTTGGCC 515  
 QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162  
 DB 516 AATGCTTCCAGCAGGAAGCCAGACCTCAGAGGAAGAACGCTTCTTCTGAGTCAGCG 575  
 QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
 DB 576 GTTCCAGCTGGGCGACACCTATGTGGATCTGGATACGAGGTGGACAAATGCCAGAAC 635  
 QY 183 LeuAspPheileHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGly 202  
 DB 636 CTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGAGAAAGTCAAGGGA 695  
 QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
 DB 696 CATACAGCCCTCTACAGAGGCAAGAGAGTGTGTGAGCAGCAGCCAGCCTCAACGTG 755  
 QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242  
 DB 756 GATGCTGTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGTATCCTTGGC 815  
 QY 243 PheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyileGlyAla 262  
 DB 816 ATGCTTACTACGAGCGCTCTTCACTGGCTCTCATCAGACACCCAGAGTGGGGCC 875  
 QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
 DB 876 CCAGCCACAGGGTCTGGCAGCTCCAGGCCCTTCAACAGGAAGGAGGATGTGCTGCTAC 935  
 QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
 DB 936 TATGAAGTGTGCTCTCTGG-----AAGGGGGCCACCAACAGAGAAATCCAGATCAGAG 989  
 QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
 DB 990 GTGCCCTACATCTCCGGGACACACAGTGGTGGGCTTTGATGATGTGAGAGCTTCAAA 1049  
 QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaileAsp 342  
 DB 1050 ACCAAGGTGAGTATCTGAAGCAGAAAGGAGTGGGGGGGGCATGTGCTGGGCACTGGAC 1109  
 QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
 DB 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGGCCGATACCCCTCATCCAGAGCTTA 1169  
 QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378  
 DB 1170 CGGCAGGAAGTGAAGTCTTCATCTTCCATCTTCCAGGCCACCCAGAGCTTGAAGTTCCAAA 1229  
 QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395  
 DB 1230 CCAGGTGAGCCCTCTGAACCT-----GAG 1253  
 QY 396 SerGlyGlySerSerGlyLysSerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415  
 DB 1254 CATGCCCCCGCCCTGGACACAGACAGCTTCTGCCAGGGCAAGCTGATGGCTCTATCCC 1313  
 QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435  
 DB 1314 AATCCTCGGGAAGCGTCCAGCTTCTACAGCTGTGCGGGGGCGGCTGTTCAGCAAGAGC 1373  
 QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451  
 DB 1374 TGCCCGCAGAGCGCTGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1421

RESULT 13

ABL57381

ID ABL57381 standard; cDNA; 1656 BP.

XX

ABL57381;  
 12-AUG-2002 (first entry)  
 Human chitinase cDNA clone MO-13B.  
 Chitinase; enzyme; human; fungicide; antifungal; infection;  
 candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;  
 paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
 Pneumocystis; gene; ss.  
 Homo sapiens.  
 Key Location/Qualifiers  
 CDS 27..1427  
 /tag= a  
 /product= "Chitinase"  
 sig\_peptide 27..89  
 /tag= b  
 mat\_peptide 90..1424  
 /tag= c  
 US6372212-B1.  
 16-APR-2002. 97US-0877599.  
 16-JUN-1997; 96US-0663618.  
 14-JUN-1996; 96US-0663618.  
 (ICOS-) ICOS CORP.  
 Gray PW;  
 WPI; 2002-442449/47.  
 P-PSDB; ABB76292.  
 Co-administering chitinase to improve the effectiveness of fungicidal  
 drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
 infections e.g. candidiasis, coccidioidomycosis and blastomycosis -  
 Example 1; Column 31-34; 26pp; English.  
 The present sequence is that of cDNA clone MO-13B (ATCC 98078)  
 encoding human chitinase (see ABB76292). The clone was isolated  
 from a cDNA library prepared from peripheral blood monocyte-derived  
 macrophages using a probe based on human chitinase clone MO-218  
 (see ABL57380). The 2 clones contain a single nucleotide  
 difference in the coding region, causing amino acid position 81 of  
 the mature protein to be either serine or glycine. Northern blots  
 showed highest chitinase gene expression in lung and ovary tissues.  
 Expression in lung is consistent with a protective role against  
 pathogenic organisms that contain chitin. The invention provides  
 human chitinase polynucleotides and polypeptides, and materials and  
 methods for the recombinant production of human chitinase products,  
 which are expected to be useful as products for treating fungal  
 infections or for the development of such products. Human  
 chitinase has a synergistic effect on the actions of other  
 fungicides. It can be administered to improve the antifungal  
 activity of a non-chitinase antifungal agent, especially  
 amphotericin B or itraconazole, in the treatment of a fungal  
 infection such as candidiasis, aspergillosis, coccidioidomycosis,  
 blastomycosis, paracoccidioidomycosis, histoplasmosis,  
 cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,  
 dermatophytoses and Pneumocystis infections (all claimed). In  
 particular, the fungal infection involves Candida, Aspergillus  
 and/or Cryptococcus spp., whose growth is not effectively  
 inhibited by contact with human chitinase alone.  
 Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;  
 Alignment Scores:

Pred. No.:	9.53e-107	Length:	1656
Score:	1295.50	Matches:	234
Percent Similarity:	66.89%	Conservative:	71
Best Local Similarity:	51.32%	Mismatches:	130
Query Match:	51.97%	Indels:	21
DB:	24	Gaps:	3

US-10-004-219B-9 (1-452) x ABL57381 (1-1656)

Qy	3	LeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro	22
Db	96	CTGGTCTGCTACTTCACCACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCTGTCGCC	155
Qy	23	AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn	42
Db	156	AAGACTTGGACCCAGCCCTTTGACCCACCTCATCTACGCTTCGCTGGCATGCCAAC	215
Qy	43	AsnGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu	62
Db	216	CACCAGCTGAGCACCCACTGAGTGAATGACGAGACTCTCTACGAGGTTCAATGGCCTG	275
Qy	63	LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr	82
Db	276	AAGAAGATGAATCCCAAGCTGAAGACCTGTAGCCATCGGAGGCTGGAATTTACGACT	335
Qy	83	AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal	102
Db	336	CAGAAGTTCACAGATATGTTAGCCAGCCCAACACCTGTGACCTTTGTCACTCGGCC	395
Qy	103	IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly	122
Db	396	ATCAGGTTTCTGCGCAATACAGCTTTTCAGCGCTTGACCTTGACTGGGAGTACCAGGA	455
Qy	123	SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg	142
Db	456	AGCCAGGGGAGCCCTGCGGTAGCAAGAGGAGGCTTTCACAACTCTGGTACGACTTGGCC	515
Qy	143	GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla	162
Db	516	AATGCTTTCAGCAGGAGAGCCAGACCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	575
Qy	163	ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr	182
Db	576	GTTCAGCTGGGAGACCTATGCTGATCTGATGATGATGATGATGATGATGATGATGATG	635
Qy	183	LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly	202
Db	636	CTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTACGGGA	695
Qy	203	GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal	222
Db	696	CATACACCCCTCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	755
Qy	223	AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly	242
Db	756	GATGCTGCTGTGCAACACTGGCTGCGAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG	815
Qy	243	PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla	262
Db	816	ATGCTACTACGAGGAGGCTCTTACACTGGCTCTCATCAGACACAGGAGTGGGGGCC	875
Qy	263	ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr	282
Db	876	CCAGCCACAGGCTTGGGACTCTCCAGGCCCCCTTCCAAAGGAGGAGGAGGAGGAGGAGGAG	935
Qy	283	TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu	302
Db	936	TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG	989
Qy	303	ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer	322
Db	990	GTGCCCTACATCTTCCGGGCAACACGATGGGTGGGCTTTGTGATGTGGAGAGCTTCAA	1049



```
db 913 ATGCCTACCTACGGACGCTCTCTACACTGGCTCTCTATCAGACACACAGATGGGGCC 972
Qy 263 ProThrSerGlyAspGlyProAlaGlyAlaThrArgGlnAlaGlyPheThrAlaTyr 282
Db 973 CAGCCACAGGCTGGCACTCAGGCCCTTCACCAAGGAAGGAGGATGTCGGCTTAC 1032
Qy 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
Db 1033 TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 1086
Qy 303 ValProTyrAlaTyrLysAlaGlnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
Db 1087 GTGCCCTACATCTCCGGACACACAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAA 1146
Qy 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
Db 1147 ACCAAGGTGAGTATCTAGACAGAGGACTGGGGGGCCATGCTCTGGGCACTGGAC 1206
Qy 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPhePheProLeuThrSerThrLeu 362
Db 1207 TTAGATGACTTTTCCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGCTA 1266
Qy 363 AsnLysAlaLeuGlyLysSerThrGluGlyCysThrAlaProAspVal----- 378
Db 1267 CGGCAGGAACAGTCTTCCATACTTGCCTTCAGGCACCCAGAGCTTGAAGTTCAAAA 1326
Qy 379 -----ProSerGluProValThrThrProProGlySerGlySerGlyGlySer 395
Db 1327 CAGGTGAGCCCTCTGAACCT-----GAG 1350
Qy 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
Db 1351 CATGGCCCGAGCCCTGGACAGACAGCTTCTGCGAGGCAAGCTGATGGGCTCTATCCC 1410
Qy 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
Db 1411 AATCTCGGGACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTGTTCCAGCAAGC 1470
Qy 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
Db 1471 TGCCGACAGGCGCTGGTGTTCAGCACTCTCTGCAATGCTGCACCTGG 1518

RESULT 15
ID AAT50834
XX AAT50834 standard; cDNA; 1713 BP.
AC AAT50834;
XX
XX 24-MAR-1997 (first entry)
XX Human chitinase cDNA clone chi.39.
DE
XX Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipidoses; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 13..1176
XX FT /*tag= a
XX FT sig_peptide 13..75
XX FT /*tag= b
XX FT mat_peptide 76..1173
XX FT /*tag= c
XX
XX WO9640940-A2.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-NL00225.
XX
PF
```

```
XX 07-JUN-1995; 95US-0486839.
XX (UNAM ) UNIV AMSTERDAM.
XX Aerts JMGF;
XX WPI; 1997-118698/11.
XX P-PSDB; AAW08585.
XX New human chitinase - used to treat or prevent infection by
XX chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
XX foods, implants etc.
XX Claim 2; Page 42-43; 58pp; English.
XX A cDNA clone (AAT50834), designated chi.39, codes for a 39 kDa human
XX chitinase (AAW08584) that is stable to many proteases, active at pH 3-
XX 8 and up to 50 deg, and stable in the circulation. Clones chi.39
XX and chi.50 (see also AAT50833) were isolated from a human macrophage
XX cDNA library using as probe a partial clone obtd. using primers
XX (see also AAT50835-36) based on a chitotriosidase purified from a
XX type 1 Gaucher disease patient. The 2 cDNA clones are the result
XX of alternative splicing of RNA. Chitinase nucleic acid can be used
XX for large-scale prodn. of recombinant human chitinases, or can be
XX incorporated into a gene therapy vector to treat or prevent
XX infection by chitin-contg. pathogens.
XX Sequence 1713 BP; 380 A; 503 C; 465 G; 365 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 7,12e-106 Length: 1713
Score: 1286.00 Matches: 240
Percent Similarity: 66.52% Conservative: 72
Best Local Similarity: 51.17% Mismatches: 134
Query Match: 18 Indels: 24
DB: 5

US-10-004-219B-9 (1-452) x AAT50834 (1-1713)
Qy 3 LeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
Db 82 CTGGTCTGCTACTTTCACCAACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTCTCGCC 141
Qy 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
Db 142 AAGGACTTGGACCCCGCCGCTTTCACCCACCTCATCTACGCTTCGCTGGCATGACCAAC 201
Qy 43 AsnGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
Db 202 CACCAGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTTACCAGGAGTTCAATGGCCTG 261
Qy 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPheGlyThr 82
Db 262 AAGAAGATGAATCCCAAGCTTGAAGACCTGTGTAGCATCGGAGGCTGGAATTTCCGCACT 321
Qy 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
Db 322 CAGAAGITTCACAGATATGTTAGCCACGCGCCACACCCCTTCAGACCTTTGTCAACTCGGCC 381
Qy 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 382 ATCAGTTTCTTCGCAATAATACAGCTTTTCACGCGCTTGACCTTGAGTGGAGTACCAGGA 441
Qy 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
Db 442 AGCCAGGGGAGCCCTGCGGTAGACCAAGGAGGCTTCAACACCTCGTACAGACTTGGCC 501
Qy 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
Db 502 AATGCTTTCAGCAGGAAGCCACACCTCAGGGAAGGAACGCTTCTTCTGAGTGCAGCG 561
Qy 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182
```

[illegible]

Search completed: June 29, 2003, 21:26:40  
Job time : 280.065 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:12:38 ; Search time 64.5237 Seconds  
(without alignments)  
2148.326 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLCYFTNMQAYRPLGSGF.....QQHCQAGLVFTSCNCCNWP 452

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlpl  
-Q=/cgn2\_1/USPTO\_spool/US10004219/runat\_24062003\_160229\_8184/app\_query.fasta\_1.1877  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MOD=LOCAL -OUTFMT=pfco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219@cgn\_1\_1\_108@runat\_24062003\_160229\_8184 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCITUS.COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result. No.	Score	Query Match	Length	ID	Description
1	1301.5	52.2	1636	US-09-039-198A-1	Sequence 1, Appli
2	1301.5	52.2	1636	US-08-877-599-1	Sequence 1, Appli
3	1301.5	52.2	1636	US-09-267-574-1	Sequence 1, Appli
4	1301.5	52.2	1643	US-08-486-839-3	Sequence 3, Appli
5	1301.5	52.2	1643	US-09-151-011-3	Sequence 3, Appli
6	1301.5	52.2	1643	US-09-343-623-3	Sequence 3, Appli
7	1295.5	52.0	1656	US-09-039-198A-3	Sequence 3, Appli
8	1295.5	52.0	1656	US-08-877-599-3	Sequence 3, Appli
9	1295.5	52.0	1656	US-09-267-574-3	Sequence 3, Appli
10	1286	51.6	1713	US-08-486-839-5	Sequence 5, Appli
11	1286	51.6	1713	US-09-151-011-5	Sequence 5, Appli
12	1286	51.6	1713	US-09-343-623-5	Sequence 5, Appli

13	971.5	39.0	1681	5	PCT-US94-07754-4	Sequence 4, Appli
14	966.5	38.8	1681	2	US-08-581-527-4	Sequence 4, Appli
15	950.5	38.1	1433	1	US-08-694-915-1	Sequence 1, Appli
16	950.5	38.1	1526	1	US-08-694-915-3	Sequence 3, Appli
17	863.5	34.6	2452	2	US-08-524-051-1	Sequence 1, Appli
18	802	32.2	1677	4	US-09-545-814-13	Sequence 13, Appli
19	802	32.2	1677	4	US-09-545-814-15	Sequence 15, Appli
20	802	32.2	1749	4	US-09-545-814-4	Sequence 4, Appli
21	802	32.2	1749	4	US-09-545-814-6	Sequence 6, Appli
22	802	32.2	2610	4	US-09-545-814-1	Sequence 1, Appli
23	802	32.2	2610	4	US-09-545-814-3	Sequence 3, Appli
24	799	32.0	1919	4	US-09-545-814-31	Sequence 31, Appli
25	729.5	29.3	1608	4	US-09-292-225-14	Sequence 20, Appli
26	729.5	29.3	1608	4	US-09-292-225-20	Sequence 22, Appli
27	729.5	29.3	1665	4	US-09-292-225-17	Sequence 17, Appli
28	729.5	29.3	1665	4	US-09-292-225-19	Sequence 19, Appli
29	729.5	29.3	1752	4	US-09-292-225-14	Sequence 14, Appli
30	729.5	29.3	1752	4	US-09-292-225-16	Sequence 16, Appli
31	729.5	29.3	1752	4	US-09-292-225-40	Sequence 40, Appli
32	722	29.0	1470	4	US-09-292-225-42	Sequence 42, Appli
33	722	29.0	1470	4	US-09-292-225-37	Sequence 37, Appli
34	722	29.0	1527	4	US-09-292-225-39	Sequence 39, Appli
35	722	29.0	1527	4	US-09-292-225-34	Sequence 34, Appli
36	722	29.0	1621	4	US-09-292-225-36	Sequence 36, Appli
37	722	29.0	1621	4	US-09-292-225-34	Sequence 34, Appli
38	667.5	26.8	1478	4	US-09-545-814-28	Sequence 28, Appli
39	667.5	26.8	1478	4	US-09-545-814-30	Sequence 30, Appli
40	470.5	18.9	1167	1	US-07-939-501A-6	Sequence 6, Appli
41	470.5	18.9	1167	4	US-08-448-398-10	Sequence 10, Appli
42	470.5	18.9	1320	1	US-07-939-501A-14	Sequence 14, Appli
43	470.5	18.9	1364	1	US-07-939-501A-13	Sequence 13, Appli
44	470.5	18.9	1405	1	US-07-939-501A-9	Sequence 9, Appli
45	461.5	18.5	1554	1	US-08-045-269C-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-039-198A-1

; Sequence 1, Application US/09039198A

; Patent No. 6200951

; GENERAL INFORMATION:

; APPLICANT: Tjoelker, Larry W.

; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 South Wacker Drive/6300 Sears Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/039.198A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Rin-Laures, Li-Hsien

; REGISTRATION NUMBER: 33,547

; REFERENCE/DOCKET NUMBER: 27866/34391

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1636 base pairs

; TYPE: nucleic acid

STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 2..1399  
 FEATURE:  
 NAME/KEY: mat peptide  
 LOCATION: 65..1399  
 US-09-039-198A-1

## Alignment Scores:

Pred. No.: 6,7e-121 Length: 1636  
 Score: 1301.50 Matches: 235  
 Percent Similarity: 67.11% Conservative: 71  
 Best Local Similarity: 51.54% Mismatches: 129  
 Query Match: 52.21% Indels: 21  
 DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-039-198A-1 (1-1636)

```

QY 3 LeuIleCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 CTGGTCTGTACTTCAACCACTGGGCCCACTACAGACAGGGGGAGGCTCGCTTCTCGCCC 130

QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 AAGGACTTGGACCCCGCCAGCTTTGCACCCCACTCATCTACGCTTCTCGCATGACCAAC 190

QY 43 AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 CACCAGCTGACGACCACTGAGTGAATGACGAGACTCTCTACAGGAGTTCAATGGCGCTG 250

QY 63 LysAsnArgAsnSerLysLeuThrLeuAlaIleGlyThrAsnPheGlyThr 82
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 AAGAAGATCAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGCTGGAATTTCCGCACT 310

QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 311 CAGAAGTTCACAGATATGTTAGCCAGCGGCCCAACACCGTCAGACCTTTGTCACTCGGCC 370

QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 371 ATCAGGTTTTCGCAAAATACAGCTTTGACGGCTTGACCTTGACCTTGAGTACCCAGGA 430

QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 431 AGCCAGGGGAGCCCTGCCCTGACACAGGAGCGCTTCAACACCTGTGTACAGACTTGGCC 490

QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 491 ANTGCCTTCACAGAGGAGCCAGACCTCAGGAGGAGGAGCGCTTCTCTGAGTGCAGCG 550

QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 551 GTTCCAGCTGGCAGACCTATCTGGATGCTGTGATCAGAGGTGGACAAAATCGCCAGAAC 610

QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 611 CTGGATTTTGTCAACCTTATGCTCCTACGACTTCCATGGCTCTTGGGAGAGTCAAGGGA 670

QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 671 CATACAGCCCTCTACAGAGGAGGAGAGAGAGTGTGACGAGCCAGCCCTCAACGTG 730

QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyValaProAlaGluLysLeuIleValGly 242
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 731 GATGCTGCTGCAACAGTGGCTGCGAGAGGAGGAGCCCTGCGCAGCAAGCTGATCTTGGC 790

QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 791 ATGCTTACCTACGAGCGCTCTTCACTGCTCTCATCATCAGACCAAGAGTGGGGGCC 850

```

```

QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 851 CCAGCCACAGGCTGTGGCACTCCAGGCCCTTCACCAAGAGGAGGATGCTGCGCTAC 910

QY 283 TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 911 TATGAAGTCTGCTCTCTGG-----AAGGGGGCCACCAACACAGAAATCCAGGATCAGAAG 964

QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 965 GTGCCCTACATCTCCGGGCAACACGAGGGGTGGCTTTGATGATGTGGAGCTTCAAA 1024

QY 323 VallysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1025 ACCAAGTTCAGTATCTGACGACAGAGGAGCTGGGGGGGCCATGCTTGGGCACTGGAC 1084

QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1085 TTAGATGACTTTGCGGCTTCTCTCTGCAACACGAGGCGCATACCCCTCATCCAGACGCTA 1144

QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1145 CGGACGAACTAGTCTTCCATACCTTGCCTTCAGGACCCAGAGCTTGAAGTCCAAA 1204

QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1205 CCAGTTCAGCCCTCTGAACCT-----GAG 1228

QY 396 SerGlyGlySerSerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1229 CATGCCCCAGCCCTGGGACAGACACAGCTTCTGCCAGGCGCAAGCTGATGGGCTCTATCCC 1288

QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1289 AATCTCGGAAACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTGTTCAGCAAAAGC 1348

QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1349 TGCCGACAGGCTGTGTCTTCCAGCAACTCTCTGCAAAATGCTGCACCTGG 1396

RESULT 2
US-08-877-599-1
; Sequence 1, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300

```

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1636 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2..1399

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 65..1399

US-08-877-599-1

## Alignment Scores:

Pred. No.: 6, 7e-121 Length: 1636

Score: 1301.50 Matches: 235

Percent Similarity: 67.11% Conservative: 71

Best Local Similarity: 51.54% Mismatches: 129

Query Match: 52.21% Indels: 21

DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-08-877-599-1 (1-1636)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
 DB 71 CTGGTCTGCTACTTCACCAACTGGGCCCGGAGTACACAGAGGGGAGGCTTCCTGGCC 130

QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42  
 DB 131 AAGGACTTGGACCCCGAGCTTTGACCCACCTCATCTACGCCCTTCGCTGGCATGACCAAC 190

QY 43 AsnGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
 DB 191 CACCAGCTGAGCACCACTGAGTGAATGACGAGACTCTTACGAGAGTTCAATGGCCTG 250

QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPheGlyThr 82  
 DB 251 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCATCGGAGGCTGGAATTCGGCACT 310

QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
 DB 311 CAGNAGTTTACAGATATGTTAGCCACGCGCCACACACCTTCAGACCTTTGTCACACTGGCC 370

QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
 DB 371 ATCAGGTTTCTCGCAAAATACACAGCTTTGACGGCCTTGACCTTGAGTGGGAGTACCGAGA 430

QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
 DB 431 AGCCAGGGGAGCCCTGCGGTAGCAAGAGGCGCTTCACACACCTCGGTACAGACTTGGCC 490

QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162  
 DB 491 ATGCTCTTCACAGGAGCCAGACCTTCAGGGAAGGAGCGCTTCTCTGAGTGCAGCG 550

QY 163 ValAlaGlyLysSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
 DB 551 GTTCAGCTGGGCAGACCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 610

QY 183 LeuAspPheIleHisValMetThrThrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
 DB 611 CTGGATTTTGTCAACCTTATGCTTACGACTTCCATGGCTCTTGGGAGAGTCAAGGGA 670

QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
 DB 671 CATACAGCCCTCTACAGAGGCGGAGAGAGAGTGGTGGAGCGGAGCTTCAAGCTG 730

QY 223 AspTyrValMetAsnTyrTrpLysAsnGlyValProAlaGluLysLeuileValGly 242

DB 731 GATGCTGCTGTGCAACAGTGGCTGCGAAGAGGGGCCCTCGCCAGCAAGCTGATCCTTGGC 790

QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262

DB 791 ATGCTACCTACGAGCGCTCTTCACTGGCTCTCTCATCAGACACAGAGTGGGGGCC 850

QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282

DB 851 CCAGCCACAGGGTCTGGCACTCCAGGCCCTTCAACAGGAAGAGGAGTGTGGCTCTAC 910

QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302

DB 911 TATGAGTCTGCTCTCTGG-----AAGGGGGCCCAACACAGAAATCCAGATCAGAA 964

QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322

DB 965 GTGCCCTACATCTTCCGGGCAACACAGTGGTGGGCTTTGATGATGGAGAGCTTCAAA 1024

QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342

DB 1025 ACCAAGGTGACGTATCTGAAGCAGAGGAGCTGGGGGGGCGCATGGTCTGGGCACCTGAC 1084

QY 343 LeuAspAspPheThrGlySerPheCysAspGlnLysPheProLeuThrSerThrLeu 362

DB 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGGCGGATACCCCTCATCCAGACGCTA 1144

QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378

DB 1145 CGGAGGAGTGTAGTCTTCCATCTTGGCTTCAGGCACCCCGAGAGCTTGAAGTTCAAA 1204

QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395

DB 1205 CCAGTTCAGCCCTCTGAACCT-----GAG 1228

QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415

DB 1229 CATGGCCCGACCCCTGGCAAGACACACGTTCTGCCAGGGCAAGCTGATGGGCTCTATCCC 1288

QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435

DB 1289 AATCTCGGGAACGGTCCAGCTTTACAGCTGTGCGAGGGGGGGCTGTTCCAGCAAAAGC 1348

QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451

DB 1349 TGCCGAGCAGGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1396

## RESULT 3

US-09-267-574-1

; Sequence 1, Application US/09267574

; Patent No. 6399571

; GENERAL INFORMATION:

; APPLICANT: Gray, Patrick W.

; APPLICANT: Tjoelker, Larry W.

; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

; FILE REFERENCE: 27866/35407

; CURRENT APPLICATION NUMBER: US/09/267,574

; CURRENT FILING DATE: 1999-03-12

; EARLIER APPLICATION NUMBER: 09/039,198

; EARLIER FILING DATE: 1998-03-12

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1636

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (2)..(1399)

; FEATURE:

; NAME/KEY: mat\_peptide

; LOCATION: (65)..(1399)

; US-09-267-574-1

## Alignment Scores:

Pred. No.: 6.7e-121 Length: 1636  
 Score: 1301.50 Matches: 235  
 Percent Similarity: 67.11% Conservative: 71  
 Best Local Similarity: 51.54% Mismatches: 129  
 Query Match: 52.21% Indels: 21  
 DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-267-574-1 (1-1636)

```

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTrpArgProGlyLeuGlySerPheLysPro 22
Db 71 CTGGTCTGTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCTGGCCC 130
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
Db 131 AAGGACTTGGACCCAGCCCTTTGACCCACCTCATCTAGCCCTTCCTGGCATGACCAAC 190
QY 43 AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
Db 191 CACCAGCTGAGCACCACTGAGTGGATGACGAGACTCTCTACAGAGTTCAATGGCCGTG 250
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
Db 251 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGCTGGAATTTGGCACT 310
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
Db 311 CAGAAGTTTACAGATATGTAGCCAGCGGCCACACCCGTGACAGCTTTGTCAACTCGGCC 370
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 371 ATCAGGTTTCTGGCAAAATACAGCTTTGACGGCTTGACCTTGACCTGGAGTACCAGGA 430
QY 123 SerArgLysSerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
Db 431 AGCCAGGGAGGCCCTCCCTAGACAAAGAGCGCTTACCAACCTGTGTACAGGCTGGCC 490
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
Db 491 ATGCTCTTCAGCAGGAAGCCAGACTCAGGAAGGAACGCTTCTTCTGAGTGAGCGG 550
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
Db 551 GTTCCAGCTGGGCAGACCTATGTGGATGTGGATACAGGTGACAAATGCCCAGAAC 610
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
Db 611 CTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAAGGTACGGGA 670
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
Db 671 CATAACAGCCCTCTTACAAAGAGGCAAGAGAGTGGTGACAGCCAGCTTCAACGTG 730
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyValaProAlaGluLysLeuIleValGly 242
Db 731 GATGCTGTGTGCAACAGTGGCTGACAGAGGGGACCCCTGCCAGCAGCTGATCTTGGC 790
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
Db 791 ATGCTTACCTACGAGGCTTCCCTTACACTGGCTCTCATCAGACACCAAGAGTGGGGCC 850
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
Db 851 CAGCCAGAGGCTTGGCACTCAGGCCCTTCCACCAAGAGAGGAGGATGTGGCCCTAC 910
QY 283 TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
Db 911 TATGAAGTGTCTCTGG-----AAGGGGGGCCACCAACAGAGAATCCAGGATCAGAG 964
QY 303 ValProTyrAlaTyrLysAlaGlnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
Db 965 GTGCCCTATCTTCGGGACCAACCAAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAA 1024

```

```

QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
Db 1025 ACCAAGGTGAGTCTTCTGAAGCAGAGGAGCTGGCGGGCCCATGGTCTGGGCATGGAC 1084
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
Db 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACACGAGGCGCGATACCCCTCATCCAGAGGCTA 1144
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
Db 1145 CGGCAGGAAGTCTGAGTCTTTCATCTTTCAGGCACCCAGAGCTTGAAGTTCCAAAA 1204
QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlyGlySer 395
Db 1205 CCAGTCTAGCCCTCTGAACCT-----GAG 1228
QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
Db 1229 CATGCCCCAGCCCTGGACAAGACACAGCTTCTGCCAGGCGCAAGCTGATGGGCTCTATCC 1288
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
Db 1289 AATCCTCGGGAACGCTCCAGCTTCTACAGCTGTGTGAGCGGGCGGCTGTTCAGCAAAAGC 1348
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
Db 1349 TGCCCGCAGGCGCTGTGTTCAGCAACTCTTGCAAAATGCTGCACCTGG 1396

```

## RESULT 4

US-08-486-839-3  
 ; Sequence 3, Application US/08486839  
 ; Patent No. 5928928  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: A human chitinase, its recombinant  
 ; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hoffmann & Baron  
 ; CITY: Jericho  
 ; STATE: New York  
 ; COUNTRY: United States of America  
 ; ZIP: 11758  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/486.839  
 ; FILING DATE: 07 - June - 1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Baron, Ronald J.  
 ; REGISTRATION NUMBER: 29,281  
 ; REFERENCE/DOCKET NUMBER: 294-26  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (516) 822-3550  
 ; TELEFAX: (516) 822-3582  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1643 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHETICAL: NO  
 ; US-08-486-839-3

Alignment Scores:

Pred. No.: 6,74e-121 Length: 1643  
Score: 1301.50 Matches: 235  
Percent Similarity: 67.11% Conservative: 71  
Best Local Similarity: 51.54% Mismatches: 129  
Query Match: 52.21% Indels: 21  
DB: 2 Gaps: 3

US-10-004-219B-9 (1-452) x US-08-486-839-3 (1-1643)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
DB 82 CTGGTCTGCTACTTACCAACTGGGCCAGTAGACACAGGGGGAGGCTCGCTTCTGCCC 141  
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42  
DB 142 AAGGACTGGACCCAGCCCTTGGACCCACCTCATCTAGCCCTTCGCTGGCATGCCAAC 201  
QY 43 AenGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
DB 202 CACCACTGAGCACCACCTGAGTGAATGACGAGACTCTCTACGAGAGTTCAATGGCCTG 261  
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82  
DB 262 AAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCATCGGAGGCTGGAAATTCGCACT 321  
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
DB 322 CAGAGTTTCACAGATATGTAGCCAGCGGCCACACACCTTGTCACTCGGCC 381  
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
DB 382 ATCAGGTTTCTGGCAAAATACAGCTTGTACGGCTTGTACCTTGTACCTGGAGTACCAAGA 441  
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
DB 442 AGCCAGGGAGGCGCTCGCTAGCAAGGAGCGCTTCAACACCTGTGTACAGACTTGGCC 501  
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162  
DB 502 ANTGCCTTCACAGGAGGAGCCAGACCTCAGGAGGAGGAGGCTTCTTGTAGTGCAGG 561  
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182  
DB 562 GTTCCAGCTGGGACAGCTATGTGGATGCTGTGATACGAGTGGACAAATCGCCAGAAC 621  
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
DB 622 CTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTACGGGA 681  
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
DB 682 CATACAGCCCTCTACAGAGGCGGAGAGAGAGTGTGACGAGCCAGCTCAACGTG 741  
QY 223 AspTyrValMetAsnTyrTrpLysAsnAenGlyAlaProAlaGluLysLeuIleValGly 242  
DB 742 GATGCTGTGTCAACAGTGGCTGCAAGAGGGGAGCCCTGCCAGCAAGCTGATCTCTTGC 801  
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
DB 802 ATGCTTACCTAGCGAGGCTCTCTTACCTACCTGGCTCTCATACGACCCAGAGTGGGGCC 861  
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
DB 862 CCAGCCAGAGGTCTGGCACTCCAGCCCTTCAACAGAGAGGAGGATGTGGCTTAC 921  
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
DB 922 TATGAAGTCTGCTCTGG-----AAGGGGGCCACCACACAGAGATCCAGATCAGAAG 975  
QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
DB 976 GTGCCCTATCTTCGGGACAAACCAAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1035

323 VallysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342  
1036 ACCAAGGTGAGTATCTGAAGCAGAGAGGACTGGGCGGGCCATGGTCTGGGCACTGGAC 1095  
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
DB 1096 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGGCCGATACCCCTCATCCAGACGCTA 1155  
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378  
DB 1156 CGCAGGAACTAGTCTTCCATCTTCCAGCCACCCAGAGCTTGAAGTTCACAAA 1215  
QY 379 -----ProSerGluProValThrThrProGlySerGlyGlyGlySer 395  
DB 1216 CCAGGTGAGCCCTCTGAACCT-----GAG 1239  
QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415  
DB 1240 CATGCCCCAGCCCTTGGACACACACAGCTTCTGCCAGGCAAGCTGATGGGCTCTATCCC 1299  
QY 416 ValAlaAspAspAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435  
DB 1300 AATCTCCGGAACGGTCCAGCTTCTACAGCTGTGCGCGGGGGCGGCTGTCAGGCAAGC 1359  
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTyr 451  
DB 1360 TGCCGACAGGCGCTGTGTCTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1407

RESULT 5  
US-09-151-011-3  
; Sequence 3, Application US/09151011  
; Patent No. 6057142  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant  
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 6900 Jericho Turnpike  
; CITY: Syosset  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11791  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/151,011  
; FILING DATE: 10 - September - 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Morris, Robert C.  
; REGISTRATION NUMBER: 42,910  
; REFERENCE/DOCKET NUMBER: 294-32 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1643 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
US-09-151-011-3

Alignment Scores: 6.74e-121 Length: 1643  
Pred. No.: 6.74e-121 Length: 1643

Score: 1301.50 Matches: 235  
Percent Similarity: 67.11% Conservative: 71  
Best Local Similarity: 51.54% Mismatches: 129  
Query Match: 52.21% Indels: 21  
DB: 3 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-151-011-3 (1-1643)

```
QY 3 LeuIleCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 82 CTGGTCTGCTACTTCAACCACTGGGCGCCAGTACAGACAGGGAGGCTCGCTTCTGGCCC 141
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
DB 142 AAGGACTTGGACCCAGCCCTTGGACCCACCTCATCTAGCCCTTCGCTGGCATGCCAAC 201
QY 43 AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
DB 202 CACCAGCTGAGCACCACTGAGTGGATGACGAGACTCTCTACCGAGGTTCAATGGCCTG 261
QY 63 LysAsnArgAsnSerLysLysThrLeuLeuAlaIleGlyClyTrpAsnPheGlyThr 82
DB 262 AAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCATCGAGGCTGGAAATTCGGCACT 321
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB 322 CAGAGTTTCAACAGATATGTTAGCCAGCGCCCAACACCGTTCAGACCTTTGTCAACTGGCC 381
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 382 ATCAGGTTTCTCGCAATAATACAGCTTTCAGCGCTTTCAGCTTTCAGCTTTCAGCTTTCAG 441
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
DB 442 AGCGGGAGGAGCCCTCGCTAGCAAGAGCGCTTCACAAACCTGTGTAGCTTTCAGCTTTCAG 501
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 502 AATGCTTTCAGCAGGAGGAGCCAGACCTCAGGGAAGGAGCGCTTCTTCTGAGTGCAGG 561
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
DB 562 GTTCCAGCTGGGAGACCTATGTGATGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAG 621
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
DB 622 CTGGATTTTGTCAACCTTATGCGCTACGACTTCCATGGCTCTTGGGAGAGTGTGAGGAG 681
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB 682 CATACAGCCCTCTACAGAGGCAAGAGAGTGTGTGAGCAGCAGCAGCTCAACCTG 741
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGly 242
DB 742 GATGCTGTGTGTGCAACAGTGTGTGCAAGAGGAGGAGCCCTGCCAGCAAGTGTATCTTGC 801
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB 802 ATGCTTACCTAGGAGGCTTCTTCACTTGGCTCTCTCATCAGACACAGAGTGTGGGGCC 861
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
DB 862 CCAGCCAGAGGTCTGGCACTCCAGGCCCTTCCACCAAGAGGAGGATGTGGCTTAC 921
QY 283 TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 922 TATGAAGTGTCTCTGG-----AAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 975
QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
DB 976 GTGCCCTCATCTTCGGGACACACAGCGGGGCTTTGATGATGTGGAGAGCTTCAAA 1035
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyAlaMetIleTrpAlaIleAsp 342
```

```
DB 1036 ACCAAGTTCAGCTATCTGAAGCAGAAGGAGTGGGGCGGGCCATGGTCTGGGCACTGCAC 1095
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB 1096 TTAGATGACTTTGGCGGCTTCTCTGCAACACAGGGCCGATACCCCTCATCCAGACGCTA 1155
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
DB 1156 CGGACGAGCACTGAGTCTTCCATCTTGGCTTCAGCACCCAGAGCTTGAAGTTCCAAA 1215
QY 379 -----ProSerGluProValThrThrProProGlySerGlyGlySer 395
DB 1216 CCAGGTCAGCCCTCTGAACCT-----GAG 1239
QY 396 SerGlyGlySerSerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
DB 1240 CATGCCCCAGCCCTTGGACAGACAGCTTCTGCCAGGCAAGCTGTATGGGCTCTATCCC 1299
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
DB 1300 AATCTCGGACGGTCCAGCTTCTACAGCTGTGAGCGGGGGCGCTGTTCCAGCAAGAC 1359
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
DB 1360 TGCCCGACAGGCGTGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1407
```

## RESULT 6

US-09-343-623-3  
Sequence 3, Application US/09343623  
Patent No. 6303118

## GENERAL INFORMATION:

APPLICANT: A human chitinase, its recombinant  
TITLE OF INVENTION: production, its use for decomposing chitin, its use  
TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann & Baron  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11758  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/343,623

FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,839  
FILING DATE: 07-June-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281  
REFERENCE/DOCKET NUMBER: 294-26

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO

US-09-343-623-3

Alignment Scores:

Pred. No.: 6,74e-121 Length: 1643  
Score: 1301.50 Match: 235  
Percent Similarity: 67.11% Conservative: 71  
Best Local Similarity: 51.54% Mismatches: 129  
Query Match: 52.21% Indels: 21  
DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-343-623-3 (1-1643)

Qy 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
Db 82 CTGGTCTCTACTTACCAACTGGGCCAGTACACAGAGGGGGAGCGCTTCGTCGCC 141

Qy 23 AspAlaLeuThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAlaGlyMetGlnAsn 42  
Db 142 AAGGACTTGGACCCAGCTTGGACCCACCTCATCTAGCCTTCGCTGGATGCCAAC 201

Qy 43 AsnGluileThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAlaGlyMetGlnAsn 62  
Db 202 CACCAGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTTACGAGGAGTCAATGGCGTG 261

Qy 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPheGlyThr 82  
Db 262 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTTGCGCACT 321

Qy 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
Db 322 CAGAAGTTACAGATATGATAGCCAGCCGCAACACCGTCAGACCTTTGTCACTCGGCC 381

Qy 103 IleLysPheLeuArgGlnTrpGlyPheAspGlyLeuAspGlyLeuAspTrpGluTrpProGly 122  
Db 382 ATCAGGTTTCTGGCAANTACAGCTTTGACGSCCTTGACCTTGACTGGGAGTACCCAGGA 441

Qy 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
Db 442 AGCCAGGGGAGCCCTGCGCTAGACAGGAGCGCTTCAACAACCCCTGGTACAGGACTGCC 501

Qy 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162  
Db 502 AATGCTCTCCAGCAGAGCCAGCCACTCAGGAAGGAGCGCTTCTTGAGTGCAGCG 561

Qy 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
Db 562 GTTCCAGCTGGGACAGCTATGTGGATGCTGGATACGAGGTGGACAAATGCCCCAGAAC 621

Qy 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluClyTyrThrGly 202  
Db 622 CTGGATTTTGTCAACCTTATGGCCTACGACTTCCATGGCTCTTGGGAGAAGTCAAGGGA 681

Qy 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
Db 682 CATAACAGCCCTCTACAGAGGCAAGAGAGTGGTGACAGCCAGCCGCTCAACGCTG 741

Qy 223 AspTyrValMetAsnTyrTrpLysAsnGlnValAlaProAlaGluLysLeuileValGly 242  
Db 742 GATGCTGTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAGCTGATCTTGGC 801

Qy 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
Db 802 ATGCCCTAGCTACGGAGCGCTCTTCACTTACCTGGCTCTCATCAGACACAGAGTGGGGCC 861

Qy 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
Db 862 CCAGCCAGAGGCTTGGGCACTCCAGGCCCTTCAACAGAGGAGGAGGATGTCGCCCTAC 921

Qy 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
Db 922 TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAGAG 975

Qy 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
Db 976 GTGCCCTACATCTCCGGGACCAACAGTGGGTGGCTTTGATGATGTGAGAGCTTCAAA 1035

Qy 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342  
Db 1036 ACCAAGGTGACGTATCTGAGCAGAGGACTGGCGGGGCCATGGTCTGGGCACTGGAC 1095

Qy 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
Db 1096 TTAGTAGCTTTGGCGGCTTCTCCTGCAACCCAGGGCGGATACCCGCTCATCCAGACGCTA 1155

Qy 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378  
Db 1156 CGGCAGGAACTGATCTTCCATACTTGCCTTACGGCAGCCCGGAGCTTGAAGTTCCAAAA 1215

Qy 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395  
Db 1216 CCAGGTGAGCCCTCTGAACCT-----GAG 1239

Qy 396 SerGlyGlySerSerGlySerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415  
Db 1240 CATGGCCCCAGCCCTGGACAGACACGTTCTGCCAGGCAAGCTGATGGGCTCTATCCC 1299

Qy 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435  
Db 1300 AATCCTGGGAAAGCTCCAGCTTCTACAGCTGTGACGGGGCGGCTTCCAGCAAGC 1359

Qy 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTyr 451  
Db 1360 TGCCCGCAGCGGCTGGTTCAGCAACTCCTCGCAATGCTGCACCTGG 1407

RESULT 7  
US-09-039-198A-3  
; Sequence 3, Application US/09039198A  
; Patent No. 6200951  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,198A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/34391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1656 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 27..1424  
; FEATURE:  
; NAME/KEY: mat\_peptide

LOCATION: 90..1424

US-09-039-198A-3

## Alignment Scores:

Pred. No.: 2,72e-120 Length: 1656  
Score: 1295.50 Matches: 234  
Percent Similarity: 66.89% Conservative: 71  
Best Local Similarity: 51.32% Mismatches: 130  
Query Match: 51.97% Indels: 21  
DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-039-198A-3 (1-1656)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
DB 96 CTGGTCTGTACTTACCAACTGGGCGCCAGTACAGACAGGGGAGGCTCGCTTCCTGCC 155  
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42  
DB 156 AAGGACTTGGACCCAGCCTTTGCACCCACCTCATCTACGCCCTTCGCTGGCATGACCAAC 215  
QY 43 AsnGluileThrThrileGlnTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
DB 216 CACCACTGAGCACCACTCAGTGGATGACGAGACTCTTACCGAGGTTCAATGGCCTG 275  
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyLysTrpAsnPheGlyThr 82  
DB 276 AGAAGATGATCCCAAGCTGAGACCTTTAGCCATCGAGGCTGGAAATTCAGCACT 335  
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheileThrSerVal 102  
DB 336 CAGAACTTCCACAGATATGTGTAGCCACCGGCCAACACCGTCAGACCTTTGTCAACTCGGCC 395  
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
DB 396 ATCAGGTTTCTGGCAATACAGCTTTGACGCCCTTGACCTTTCAGTGGAGTACCAGGA 455  
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuLysGluMetArg 142  
DB 456 AGCAGGGAGGCGCTCCGTAGACAGGAGCGCTTCAACACCTGGTACAGACTTGGCC 515  
QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162  
DB 516 AATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAACGCTTCTCTGAGTGCAGCG 575  
QY 163 ValAlaGlyLysSerAsnleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
DB 576 GTTCCAGCTGGCGAGACCTATGTGGATGTGTGATACGAGGTGGACAAATCGCCAGAAC 635  
QY 183 LeuAspPheileHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
DB 636 CTGGATTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAAGGTACAGGA 695  
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
DB 696 CATAACAGCCCTCTTACAAGAGGCAAGAGAGTGTGTCAGCAGCAGCCTCAACGTG 755  
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242  
DB 756 GATGCTGTGTCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATCTCTGGC 815  
QY 243 PheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyileGlyAla 262  
DB 816 ATGCTTACTAGGAGCGCTCTTCACTGGCTCTCATCAGACACCCAGAGTGGGGCC 875  
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
DB 876 CCAGCCACAGGCTTGGCACTCCAGGCCCTTCCACCAAGGAAGGAGGATGCTGGCTAC 935  
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
DB 936 TATGAAGTCTGTCTCTGG-----AAGGGGGCCACCAACACAGAGAATCCAGGATCAGAG 989

QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
DB 990 GTGGCTTACATCTCCGGGACAAACAGTGGTGGCTTGTATGATGGAGAGCTTCAA 1049  
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaileAsp 342  
DB 1050 ACCAAGTTCAGCTATCTGAAGCAGAAGGAGTGGGCGGGCCATGCTTGGGCACTGGAC 1109  
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
DB 1110 TTAGATGATCTTGGCGGCTTCTCTGCAACAGGCGCATACCCCTCATCCAGACGCTA 1169  
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378  
DB 1170 CGGACGAGTACGAGTCTTCCATCTTCCAGCCACCCAGAGCTTGAAGTTCACAAA 1229  
QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395  
DB 1230 CCAGGTCAGCCCTCTGAACCT-----GAG 1253  
QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415  
DB 1254 CATGCCCCAGCCCTGGACACACAGCTTCTGCCAGGCAAGCTGATGGGCTCTATCCC 1313  
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435  
DB 1314 AATCTCGGGAACGCTCCAGCTTCTACAGCTGTGCGGCGGGCGGCTGTTCCAGCAAGC 1373  
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451  
DB 1374 TGCCCGCAGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1421

RESULT 8  
US-08-877-599-3  
; Sequence 3, Application US/08877599  
; Patent No. 6372212  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; TITLE OF INVENTION: Chitinase Materials and Methods  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/877,599  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/663,618  
; FILING DATE: 14-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/33994  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1656 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear



MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 27..1424  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 90..1424  
US-08-877-599-3

## Alignment Scores:

Pred. No.: 2,72e-120 Length: 1656  
Score: 1295.50 Matches: 234  
Percent Similarity: 66.89% Conservative: 71  
Best Local Similarity: 51.32% Mismatches: 130  
Query Match: 51.97% Indels: 21  
DB: 4 Gaps: 3

US-10-004-219b-9 (1-452) x US-08-877-599-3 (1-1656)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
DB 96 CTGGTCTGCTACTTCCACCACTGGGCCAGTACACAGGCGGAGGCTCGCTTCTGCCC 155  
QY 23 AspAlaPheAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42  
DB 156 AAGGACTTGGACCCAGCCCTTTCACCCACCTCATCTAGCCCTTCGCTGGCATGACCAAC 215  
QY 43 AsnGluileThrIleGlnTrpAsnValThrLeuTyrLysAlaPheAsnAspLeu 62  
DB 216 CACCGCTGAGCACCAGTGGTGAATGACGAGCTCTTACCGAGGTTCATGGCCGTG 275  
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyLysTrpAsnPheGlyThr 82  
DB 276 AAGAGATCAATCCCAAGCTGAGACCTGTGTAGCCATCGAGGCTGGAATTCAGCACT 335  
QY 83 AlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
DB 336 CAGAAAGTTCACAGATATGTTAGCCAGCGCAACACCGTCAGACCTTTGTCAACTCGGCC 395  
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGlnTyrProGly 122  
DB 396 ATCAGGTTTCTCGCAATACAGCTTTGACGGCTTGTACCTTGTGAGTGGAGTACCAGGA 455  
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
DB 456 AGCCAGGGAGGAGCCCTCCCTACAGAGGAGCCCTTCACACCTGTGACAGACTTGCCC 515  
QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162  
DB 516 AATGCTTCCAGCAGAGAGCCAGACCTCAGGGAAGGAACGCTTCTTCTGAGTGCAGCG 575  
QY 163 ValAlaGlyGlyLysSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
DB 576 GTTCCAGCTGGGACACCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 635  
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGlyTyrThrGly 202  
DB 636 CTGGATTTTGTCAACCTTAGCCCTACGACTTCCATGCTTCTTGGAGAGTCAAGGA 695  
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
DB 696 CATAACAGCCCTCTTACAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCCTCAACGTG 755  
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242  
DB 756 GATGCTGTGTCACAGTGGCTGTCAGAGGAGGAGCCCTTCCAGCAAGCTGATCTTGGC 815  
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
DB 816 ATGCTTACTACGAGGCTTCTTACACTGGCTCTCTATCAGACCCAGAGTGGGGGCC 875  
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282

DB 876 CCAGCCACAGGGTCTGGCACTCCAGGCCCTTCCACAGGAAGGAGGATGCTGGCCTAC 935  
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
DB 936 TATGAAGTCTGCTCTCTGG-----AAGGGGCCACCAAAACAGAGAATCCAGATCAGAAG 989  
QY 303 ValProTyrAlaTyrLysAlaAsnGlnTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
DB 990 GTGCCCTACATCTTCCGGGCAACACAGTGGGTGGCTTTGATGATGTTGAGAGCTTCAAA 1049  
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342  
DB 1050 ACCAAGGTCACTATCTGAAGCAGAGGAGTGGGGCGGGCCATGGTCTGGGCACCTGGAC 1109  
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
DB 1110 TTATGATGCTTTGGGGCTTCTCTGCAACAGGCGGCGATACCCCTCATCCAGACGCTA 1169  
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378  
DB 1170 CGGAGGAAGTCACTGAGTCTTCCATACCTTTCAGGCACCCAGAGCTTGAAGTTCAAAA 1229  
QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395  
DB 1230 CCAGTCAAGCCCTCTGAACCT-----GAG 1253  
QY 396 SerGlyGlySerSerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415  
DB 1254 CATGCCCCAGCCCTGGACAGACAGACGTTCTGCCAGGCAAGCTGTATGGCTCTATCCC 1313  
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435  
DB 1314 AATCTCGGGAACGTCACAGCTTCTACAGCTGTGCGGGGGCGGCTGTTCAGCAAAAGC 1373  
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451  
DB 1374 TGCCGACAGGCGCTGGTGTTCAGCAACTCTCTGCAAAATGTCACCTGG 1421

## RESULT 9

US-09-267-574-3  
; Sequence 3, Application US/09267574  
; Patent No. 6399571  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; FILE REFERENCE: 27866/35407  
; CURRENT APPLICATION NUMBER: US/09/267,574  
; CURRENT FILING DATE: 1999-03-12  
; EARLIER APPLICATION NUMBER: 09/039,198  
; EARLIER FILING DATE: 1998-03-12  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1656  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (27)..(1424)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (90)..(1424)  
US-09-267-574-3

## Alignment Scores:

Pred. No.: 2,72e-120 Length: 1656  
Score: 1295.50 Matches: 234  
Percent Similarity: 66.89% Conservative: 71  
Best Local Similarity: 51.32% Mismatches: 130  
Query Match: 51.97% Indels: 21  
DB: 4 Gaps: 3



QY 3 LeuIleCysTyrPheThrAsnTrpAlaGlnTrpArgProGlyLeuGlySerPheLysPro 22  
DB 82 CTGGTCTGCTACTTACCACTGGCCAGTACAGACAGGGGGAGGCTCGCTTCCTGCC 141  
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42  
DB 142 AAGGACTTGGACCCAGCCCTTTGCAACCCACCTCATCTACGCCCTCGCTGGCATGACCAAC 201  
QY 43 AsnGluIleThrIleGlnTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
DB 202 CACCACTGAGCAGCAGTGGAGTGAATACGAGACTCTTACGAGAGTCAATGGCCGTG 261  
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr 82  
DB 262 AGAAGATGAATCCCACTGAGAGCCCTGTAGCCATCGAGGCTGGAAATTCGGCACT 321  
QY 83 AlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
DB 322 CAGAAAGTTTACAGATATGTTAGTACCGGCCAACACCGTCAACCTTTGTCAACTCGGCC 381  
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
DB 382 ATCAGGTTTCTCGCAATATACAGCTTTGACGGCCCTTGACTTGGAGTACCGAGA 441  
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
DB 442 AGCCAGGGAGCCCTGCGGTAGCAAGAGGCCCTTCAACACCTGGTACAGACTGGCC 501  
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162  
DB 502 AATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAACGCTTCTTCTGAGTGCAGCG 561  
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182  
DB 562 GTTCCAGCTGGCCAGCCTATGTGATGCTGATACGAGTGGAGTGGCAAAATCCCGCAGAAC 621  
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
DB 622 CTGGATTTGTCAACCTTATGCGCTACGATCTCCATGGCTTGGAGAGTCAAGCGGA 681  
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
DB 682 CATAACAGCCCTCTACAAGAGGCAAGAGAGTGTGCGAGCAGCCAGCCTCAACGTG 741  
QY 223 AspTyrValMetAsnTyrTrpLysAsnGlnAlaProAlaGluLysLeuIleValGly 242  
DB 742 GATGCTGTGTGCAACAGTGGCTGCAGAGAGGGGACCCCTGCGCAGCAAGCTGATCTTGGC 801  
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
DB 802 ATGCTTACTACGAGCGCTCTTCACTTGGCTCTCATCAGACACAGAGTGGGGCC 861  
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
DB 862 CCAGCCAGAGGCTTGGCACTCCAGGCCCTTCAACCAAGAGGAGGATGCTGGCTTAC 921  
QY 283 TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
DB 922 TATGAAGTGTGCTCTGG-----AAGGGGGGCCCAACACAGAGAATCCAGGATCAGAG 975  
QY 303 ValProTyrAlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSer 322  
DB 976 GTGCCCTACATCTCCGGGACACACAGTGGGGCTTTGATGATGTTGGAGAGCTTCAAA 1035  
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyValaMetIleTrpAlaIleAsp 342  
DB 1036 ACCAAGTTCAGCTATCTGAAGCAGAGAGGAGTGGGGGCCCATGCTCTGGGCACTGGAC 1095  
QY 343 LeuAspAspPheThrCysPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
DB 1096 TTAGATGACTTTTGGCGGCTTCTCTCTGCAACAGCGGCCGATATCCCTCATCCAGAGCTTA 1155

QY 363 AsnLysAlaLeu-----GlyIleSer 369  
DB 1156 CGGCAGGAACCTGAATGGGTAAAGCCTTAACTGCCTGTACATGTGAGGCCAGGTGTGGC 1215  
QY 370 ThrGluGlyCys---ThrAlaProAspValProSerGluProValThrPro----- 386  
DB 1216 TGTGG-CACTGTGCTTACGTGTAGTCTTCCATATCTTGCCTTCAGGCACCCAGAGCTT 1274  
QY 387 -----ProGlySerGlySerGlyGlyGlySerSerGlySerGlyGly 402  
DB 1275 GAAGTTCCAAACACAGTCCAGCTCT---GAACCTGAGCATGGCCCAAGCCCTGGCAA 1331  
QY 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAla 422  
DB 1332 GACAGCTTCTCCAGGGCAAGCTGATGGCTCTATCCCAATCTCTCGGGAACGGTCCAGC 1391  
QY 423 PheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442  
DB 1392 TTCTACAGCTGTGAGCGGGCGGCTGTTCAGCAAAAGCTGCCGACAGGCTGTGTGTTC 1451  
QY 443 AspThrSerCysAsnCysAsnTyr 451  
DB 1452 AGCAACTCTGCAAAATGCTGCACCTGG 1478  
RESULT 11  
US-09-151-011-5  
; Sequence 5, Application US/09151011  
; Patent No. 6057142  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant  
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in  
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; STREET: 6900 Jericho Turnpike  
; CITY: Syosset  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11791  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/151,011  
; FILING DATE: 10 - September - 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Morris, Robert C.  
; REGISTRATION NUMBER: 42,910  
; REFERENCE/DOCKET NUMBER: 294-32 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1713 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
US-09-151-011-5  
Alignment Scores:  
Pred. No.: 2,576-119 Length: 1713  
Score: 1286.00 Matches: 240  
Percent Similarity: 66.52% Conservative: 72  
Best Local Similarity: 51.17% Mismatches: 134  
Query Match: 51.58% Indels: 24

DB: 3 Gaps: 5  
US-10-004-219B-9 (1-452) x US-09-151-011-5 (1-1713)  
QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLeuPro 22  
DB 82 CTGGTCTGTACTTCCAACTGGGCGCCAGTACAGACAGGGGAGGCTCGCTTCTCTGCC 141  
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetGlnAsn 42  
DB 142 AAGGACTTGGACCCAGCTTTGCACCCACCTCATCTAGCCCTTCCTGGCATGACCAAC 201  
QY 43 AsnGluileThrThrileGluTrpAsnAspValThrLeuTyrIlysAlaPheAsnAspLeu 62  
DB 202 CACCAGCTGAGCACCACTGAGTGGATGACGAGACTCTCTACAGGAGTTCAATGGCCTG 261  
QY 63 LysAsnArgAsnSerIlysLeuThrLeuLeuAlaileGlyThrAsnPheGlyThr 82  
DB 262 AAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGGAGGCTTGAATTTGGCACT 321  
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
DB 322 CAGAAGTTACAGATATGTTAGCCAGCGGCAACACCGTCAGACCTTTGTCAACTCGGCC 381  
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
DB 382 ATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGAGTGGGAGTACCCAGGA 441  
QY 123 SerArgLysSerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
DB 442 AGCCAGGGAGGCGCTCCGCTAGACAAGAGCGCTTCCACCAACCTGTGTCAGGACTTGGCC 501  
QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162  
DB 502 AATGCTTTCAGCAGGAGCCAGACTCAGGAAGAAAGCGCTTCTCTGAGTGGCAGG 561  
QY 163 ValAlaGlyGlySerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
DB 562 GTTCCAGCTGGGAGAGCTATGTTGGATGCTGGATACGAGTGGACAAAATCGCCAGAAC 621  
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
DB 622 CTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGAGAGGTACGGGA 681  
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
DB 682 CATACAGGCCCTCTACAGAGCGCAAGAGAGTGGTGCAGCAGCCAGCTCAACGTTG 741  
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242  
DB 742 GATGCTGTGTGCAACAGTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATCCTTGGC 801  
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
DB 802 ATGCTTACCTACGAGGCTCTTCCACCTGGCTCTCATCAGACACCAAGAGTGGGGGCC 861  
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
DB 862 CCAGCCACAGGCTGTGGCACTCCAGGCCCTTCCACAGAGAGGAGGATGCTGGCTTAC 921  
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
DB 922 TATGAAGTGTCTCTGG-----AAGGGGGCCCAACACAGAGATCCAGGATCAGAG 975  
QY 303 ValProTyrAlaTyrIlysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
DB 976 GTGCCCTATCTTCGGGACCAACAGTGGGCTTTGATGATGAGTGGAGAGCTTCAA 1035  
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyValaMetIleTrpAlaIleAsp 342  
DB 1036 ACCAAGGTGAGTATCTGAAGCAGAGAGGAGTGGGGGGGCGGATGCTCTGGGACTGGAC 1095  
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362

DB 1096 TTAGATGACTTTCGCGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1155  
QY 363 AsnLysAlaLeu-----GlyIleSer 369  
DB 1156 CGGACGAGAACTAATGAGTAAAGCCCTTAAGTCCCTGTCATGTGAGGCCAGGTGTGCT 1215  
QY 370 ThrGluGlyCys---ThrAlaProAspValProSerGluProValThrPro----- 386  
DB 1216 TGTGG-CACTGTGCTTACGCTTAGTCTTCCATCTTCCACTTTCAGGACCCAGGCTT 1274  
QY 387 -----ProGlySerGlySerGlyGlyGlySerSerGlySerGlyGly 402  
DB 1275 GAAGTTTCCAAAACAGGTGAGCCCTCT---GAACCTGAGCATGGCCCAAGCCCTGGCAA 1331  
QY 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspArgAsnAla 422  
DB 1332 GACAGCTTCTCCAGGCGGCAAGCTGATGGCTCTATCCCAATCTCTCGGGAACGGTCCAGC 1391  
QY 423 PheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442  
DB 1392 TTCTACAGCTGTGACGGGGGGGGCTGTTCAGCAAAAGCTGCCCAGACGGCTGTGTTTC 1451  
QY 443 AspThrSerCysAsnCysCysAsnTrp 451  
DB 1452 AGCAACTCTGCAAAATGCTGCACCTGG 1478

## RESULT 12

US-09-343-623-5  
; Sequence 5, Application US/09343623  
; Patent No. 6303118

## GENERAL INFORMATION:

APPLICANT: A human chitinase, its recombinant  
TITLE OF INVENTION: production, its use for decomposing chitin, its use  
TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Baron  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11758

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
SOFTWARE:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/343,623

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/486,839

FILING DATE: 07-June-1995

ATTORNEY/AGENT INFORMATION:

NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1713 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

US-09-343-623-5

## Alignment Scores:

Pred. No.: 2.57e-119 Length: 1713  
Score: 1286.00 Matches: 240  
Percent Similarity: 66.52% Conservative: 72  
Best Local Similarity: 51.17% Mismatches: 134  
Query Match: 51.58% Indels: 24  
DB: 4 Gaps: 5

US-10-004-219B-9 (1-452) x US-09-343-623-5 (1-1713)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheIlePro 22  
DB 82 CTGGTCTCTACTTACCAACTGGGCCAGTACACAGAGGGAGGCTCGCTTCGCGCC 141

QY 23 AspaPileAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42  
DB 142 AAGGACTTGGACCCAGCCTTGGACCCACCTCATCTAGGCTTCGCTGGATGACCAAC 201

QY 43 AsnGluileThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62  
DB 202 CACCAGCTGAGCACCACCTGAGTGAATGACGAGACTCTTACCAAGAGTTCAATGGCCTG 261

QY 53 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82  
DB 262 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTCGGCACT 321

QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102  
DB 322 CAGAATTCACAGATATGTTAGCCAGGCCAACACCGTCAGACCTTTGTCMACTCGGCC 381

QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
DB 382 ATCAGGTTTCTGGCAATACAGCTTTGAGCGCTTGACCTTGTGAGGAGTACCGAGA 441

QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142  
DB 442 AGCCAGGGGAGCCCTGCGCTAGACAGGAGCGCTTTCACAAACCTCGTACAGACTTGCC 501

QY 143 GluAlaPheGluGlnAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162  
DB 502 AATGCTTCCAGCAGAGCCAGACCTCAGGGAAGGAGCGCTTCTCTGAGTGACGG 561

QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182  
DB 562 GTTCCAGCTGGCGACACCTATGTGATGCTGATAGAGGTGGACAAATCCGCCAGAC 621

QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202  
DB 622 CTGGATTTTGTCAACCTTATGGCCTACGACTTCCATGCTCTTGGGAGAAGTCAAGGA 681

QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222  
DB 682 CATAACAGCCCTCTACAGGCGCAGAGAGAGTGGTGGAGCAGCAGCTCAACGTG 741

QY 223 AsptyrValMetAsnTyrTrpLysAsnGlnAlaProAlaGluLysLeuileValGly 242  
DB 742 GATGCTGTGTCAACAGTGGCTGCAGAGGGGAGCCCTCGCCAGCAAGTATCCTTGGC 801

QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262  
DB 802 ATGCGCTAGCTAGGAGCGCTCTTCACTACCTGGCCCTCTCATCAGACACCAAGTGGGGCC 861

QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282  
DB 862 CAGCCACAGGCTCGGACCTCCAGGCCCTTCCACAGGAGGAGGAGGATGTCGCCCTAC 921

QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302  
DB 922 TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAGAG 975

QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322  
DB 976 GTGCCCTACATCTCCGGGACAAACAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAA 1035

QY 323 VallysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342  
DB 1036 ACCAGGTGACCTATCTGAACAGAGGACTGGGGGGGCGCATGGTCTGGGCACTGAC 1095

QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362  
DB 1096 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1155

QY 363 AsnLysAlaLeu-----GlyIleSer 369  
DB 1156 CGGAGGAACTGAATGGGTAAAGCTTAACTGCTGTGCATGTGAGGCCAGGTGTGCC 1215

QY 370 ThrGluGlyCys---ThrAlaProAspValProSerGluProValThrPro----- 386  
DB 1216 TGTGG-CACTGTGCTTCAGCTGTAGGTCTTCATACCTTTCCTTCAGGACCCAGAGCTT 1274

QY 387 -----ProGlySerGlySerGlyGlyGlySerGlySerGlySerGlyGly 402  
DB 1275 GAAGTTCCAAACCCAGGTTCAGCCCTCT---GAACTGAGCATGGCCCAAGCCTCGCAA 1331

QY 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAla 422  
DB 1332 GACAGTTCTCCAGGCGCAAGCTGTATGGCTCTATCCCAATCTCTCGGGAACGGTCCAGC 1391

QY 423 PheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442  
DB 1392 TTCTACAGCTGTGCGCGGGCGGCTGTTCAGCAGCAAGCTGCCCGACAGGCTGTGTGTT 1451

QY 443 AspThrSerCysAsnCysCysAsnTrp 451  
DB 1452 AGCAACTCTGCAAAATGCTGCACCTGG 1478

## RESULT 13

PCT-US94-07754-4  
; Sequence 4, Application PC/TUS9407754  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR  
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 90067  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/07754  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOWELLS, STACY L.  
; REGISTRATION NUMBER: 34,842  
; REFERENCE/DOCKET NUMBER: FD 3665  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/455-5100  
; TELEFAX: 619/455-5110  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1681 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; IMMEDIATE SOURCE:  
; CLONE: YKL-40

US-08-581-527-4



```
Query Match: 38.13% Indels: 13
DB: 1 Gaps: 6
US-10-004-219B-9 (1-452) x US-08-694-915-1 (1-1433)

QY 1 TyrAsnLeuLeuCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
DB 115 TACAAATCGTTGCTACTTTACCAACTGGTCCAGGACCGGAGAAACCCAGGAAAATTC 174
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 175 ACCCTTGAGAAATATGACCCCTTCCTGCTCATCTCATCTATTCATTTCCTGCGCAGCATC 234
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
DB 235 GAAACACACAGGTTATCATCAGGACAGAGTGAAGTGTCTTACAGACCATCAAC 294
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 295 AGTCTCAAAACCAAGAATCCAAACTGMAAATTCCTTGTCCATTGGAGGTACCTGTTT 354
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
DB 355 GGTTCCAAAGGGTTCACCCCTATGGTGGATTCCTTACATCACGCTTGGAAATTCATTAAAC 414
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 415 TCCATATCTCTGTTCTGAGAACCATTAACCTTTGATGGACTGGATGAAGTGGATCTAC 474
QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
DB 475 CCAGATCAGAAA-----GAAACACATCAT---TTCACCTGTGCTGATTTCATGAG 519
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
DB 520 TTAGCAGAGACCTTTTCCAGAGGACTTCCAAATAATCCAAAGGAAAGGCTTCTTTGACT 579
QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
DB 580 GCGGGGTATCTCAGGGAGGCAATGATTTGATAACAGCTATCAAGTTGAGAACTGGCA 639
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGly--- 199
DB 640 AAAGATCTGGATTTTCATCAACCTCTCTGCTTGTGACTTCCATGGGTCTTGGGAAAGCCC 699
QY 200 ---TyrThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAla 218
DB 700 CTTATCACTGGCCACACAGCCCTCTGAGCAGGGGTGGCAGGAGGCGGCAAGCTCC 759
QY 219 TyrLeuAsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLys 238
DB 760 TACTACATGTGGAATATGCTGTGGGGTACTGGATACATAAGGGAATGCCATCAGAGAAG 819
QY 239 LeuIleValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsn 258
DB 820 GTGGTCATGGGCATCCCCACATATGGGCATCTTCCATCAGCTCTCT---GCAGAAACC 876
QY 259 GlyIleGlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGly 278
DB 877 ACCGTGGGGGGCCCTGCCTCTGGCCCTGGAGCTGGAGCCCATCAGAGTCTTCAGGC 936
QY 279 PheTrpAlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAsp 298
DB 937 TTCCTGGCCCTATTATGAGATCTGCCAGTTCCTGAAA---GGAGCCAAAGATCAGCGGCTC 993
QY 299 AlaSerGlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIle 318
DB 994 CAGGATCAGCAGGTTCCTTACGAGTCAAGGGGAACCAAGTGGGTGGCTATGATGATGTG 1053
QY 319 LysSerPheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIle 338
DB 1054 AAGATGTGAGACCAAGGTTCAGTTCTTAAAGAAATTTTAAACCTGGGAGGAGCCATGATC 1113
```

Search completed: June 30, 2003, 01:11:17  
Job time : 95.5237 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 23:54:39 ; Search time 164.82 Seconds  
(without alignments)  
4070.941 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLICYFTWQAQYRPLGSGF.....QQHCOAGLVFTSCNCCNWP 452

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=slp  
-Q/cgn2\_1/USPTO\_spool/US10004219/runat\_24062003\_160230\_8285/app.query.fasta\_1.1877  
-DB=Published Applications NA -QWTF=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10004219 @CGN 1.1 123 @runat\_24062003\_160230\_8285  
-NCPUL=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCk=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*  
4: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
6: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
7: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*  
8: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
11: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*  
12: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2493	100.0	1525	9	US-10-004-219B-3
2	2089.5	83.8	1625	9	US-10-004-219B-2
3	992.5	39.8	1925	9	US-10-097-340-44
4	987.5	39.6	1038	9	US-10-202-436A-14

5	966.5	38.8	1681	10	US-09-215-077A-4	Sequence 4, Appli
6	966.5	38.8	1681	10	US-09-262-213A-4	Sequence 4, Appli
7	902.5	36.2	1391	10	US-09-822-830A-402	Sequence 402, App
8	854.5	34.3	1474	10	US-09-765-231A-8	Sequence 8, Appli
9	729.5	29.3	1608	9	US-10-218-743-20	Sequence 20, Appli
c 10	729.5	29.3	1608	9	US-10-218-743-22	Sequence 22, Appli
c 11	729.5	29.3	1665	9	US-10-218-743-17	Sequence 17, Appli
c 12	729.5	29.3	1665	9	US-10-218-743-19	Sequence 19, Appli
c 13	729.5	29.3	1752	9	US-10-218-743-14	Sequence 14, Appli
c 14	729.5	29.3	1752	9	US-10-218-743-16	Sequence 16, Appli
c 15	722	29.0	1470	9	US-10-218-743-40	Sequence 40, Appli
c 16	722	29.0	1470	9	US-10-218-743-42	Sequence 42, Appli
c 17	722	29.0	1527	9	US-10-218-743-37	Sequence 37, Appli
c 18	722	29.0	1527	9	US-10-218-743-39	Sequence 39, Appli
c 19	722	29.0	1621	9	US-10-218-743-34	Sequence 34, Appli
c 20	722	29.0	1621	9	US-10-218-743-36	Sequence 36, Appli
c 21	638	25.6	449	10	US-09-960-352-678	Sequence 678, App
c 22	476	19.1	1271	9	US-09-923-844B-1	Sequence 1, Appli
c 23	439.5	17.6	1797	10	US-09-938-842A-1684	Sequence 1684, Ap
c 24	405.5	16.3	415	10	US-09-960-352-2589	Sequence 2589, Ap
c 25	402.5	16.1	410	10	US-09-960-352-2589	Sequence 2589, Ap
c 26	402.5	16.1	410	10	US-09-960-352-2589	Sequence 2589, Ap
c 27	402.5	16.1	414	10	US-09-960-352-13050	Sequence 13050, A
c 28	402.5	16.1	424	10	US-09-960-352-3057	Sequence 3057, Ap
c 29	396.5	15.9	418	10	US-09-960-352-3072	Sequence 3072, Ap
c 30	390	15.6	399	10	US-09-960-352-265	Sequence 265, App
c 31	383	15.4	398	10	US-09-960-352-14460	Sequence 14460, A
c 32	381.5	15.3	411	10	US-09-960-352-14919	Sequence 14919, A
c 33	372.5	14.9	384	10	US-09-960-352-5721	Sequence 5721, Ap
c 34	346	13.9	2103	10	US-09-974-300-661	Sequence 661, App
c 35	323	13.0	438	10	US-09-960-352-3315	Sequence 3315, Ap
c 36	319.5	12.8	427	10	US-09-960-352-1882	Sequence 1882, Ap
c 37	319	12.8	422	10	US-09-960-352-5957	Sequence 5957, Ap
c 38	319	12.8	423	10	US-09-960-352-12100	Sequence 12100, A
c 39	316.5	12.7	437	9	US-09-918-995-35463	Sequence 35463, A
c 40	315.5	12.7	427	10	US-09-960-352-4978	Sequence 4978, Ap
c 41	315	12.6	430	10	US-09-960-352-1189	Sequence 1189, Ap
c 42	314	12.6	422	10	US-09-960-352-10642	Sequence 10642, A
c 43	313	12.6	419	10	US-09-960-352-10172	Sequence 10172, A
c 44	312.5	12.5	426	10	US-09-960-352-4979	Sequence 4979, Ap
c 45	311	12.5	419	10	US-09-960-352-4053	Sequence 4053, Ap

#### ALIGNMENTS

RESULT 1  
US-10-004-219B-3  
; Sequence 3, Application US/10004219B  
; Publication No. US20030087414A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; FILE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1525  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: mouse  
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid  
; OTHER INFORMATION: sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase  
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence  
; FEATURE:

NAME/KEY: CDS  
LOCATION: (1) (1419)  
US-10-004-219B-3

Alignment Scores:  
Pred. No.: 3.6e-281 Length: 1525  
Score: 2493.00 Matches: 452  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-004-219B-9 (1-452) x US-10-004-219B-3 (1-1525)

QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20  
DB 64 TACAATCTGATATGCTATTTACCAACTATGGCCCGCATGCGGCGAGCTTC 123  
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
DB 124 AAGCCTGATGACATTAAACCCCTGCTGTGTACTCAGCTGATCTATGCTTGTGGGATG 183  
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrIleAlaPheAsn 60  
DB 184 CAGAACAAATGAGATCACCACATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTCAAT 243  
QY 61 AspLeuIleAsnArgAsnSerIleLysLeuIleThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
DB 244 GACTTGAAAACAGGAACAGCAACCTGAAACCTCTGCGCAATTGGAGGCTGGAACCTTT 303  
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100  
DB 304 GGAACCTGCTCTTCACTACCATGGTTTCCACTCTTCAGAACCCGACAGCTTCATTACC 363  
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120  
DB 364 TCAGTCATCAAAATTTCTGCGTCAGTATGGTTTGTAGGAGCTGGACCTGGAGTGGAAATAC 423  
QY 121 ProGlySerArgGlySerProGlnAsnIleLysLeuPheThrValLeuValIleGlu 140  
DB 424 CAGGCTCAGTGGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGCTCTGGTGAAGGAA 483  
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160  
DB 484 ATGCGTGAGCTTTTTCAGCAGGAGGCTATTGAGAGCAACAGGCCAGAGCTGATGTTACT 543  
QY 161 AlaAlaValaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
DB 544 GCTGCTGATGCTGGGATTTCCAAATCCAGGCTGGCTATGAGATCCCTGAACTTTCT 603  
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerThrGluGlyTyr 200  
DB 604 AAGTACTGGATTTTCACTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663  
QY 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrGluThrGlySerAsnAlaTyrLeu 220  
DB 664 ACTGGGAGATAGTCTCTTTACAAATACCTTACTGAGCTGTGTAGCAATGCCTACCTC 723  
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240  
DB 724 AATGTGGATTATGTCATGAATATTGGAAGAAATGAGAGCCCGAGCTGAGAAGCTCAT 783  
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260  
DB 784 GTTGGATTTCCAGAGTATGACACACTTCACTCCTGAGAAACCCCTCTGATATGGAAT 843  
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280  
DB 844 GGTGCCCTTACCTCTGATGGCCCTGTGGCCCTATACAGAGAGAGCTGGTCTGG 903  
QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300  
DB 904 GCCTACTATGAGATTTGCACCTTTCTGAGAAGTGGAGCCACTGAGCTCTGGATGCTCC 963

QY 301 GlnGluValProTyrAlaTyrIleAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320  
DB 964 CAGAAGTGGCCCTATGCTATTAAGCCCAACGAGTGGCTTGGCTATGACATATCAAGAGC 1023  
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340  
DB 1024 TTCAGTGTAAAGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCC 1083  
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360  
DB 1084 ATTGACCTTGATGACTTCACCTGGCTCTTCTGTGATCAGGGGAAAATTTCTCTGACTTCT 1143  
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380  
DB 1144 ACTTTGAACAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCTCCAGCTGCTTCC 1203  
QY 381 GluProValThrThrProGlySerGlySerGlyGlySerSerGlyGlySerSer 400  
DB 1204 GAGCCAGTGACTACTCTCCAGGAAGTGGAGTGGGGTGGAAAGCTCCGGAGGAAGCTCT 1263  
QY 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420  
DB 1264 GGAGCAGTGGATTCTGTGCGCAACAAAGCAGATGGCTCTACCTGTGGCAGATGACAGA 1323  
QY 421 AsnAlaPheThrGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440  
DB 1324 AATGCTTTTGGCAGTGCATCAATGGAATCATACACAGCAGCATTTGTCAGAGGGCTT 1383  
QY 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452  
DB 1384 GTTTTGTATACAGCTGTAATTGCTGCAACTGGCCA 1419

# RESULT 2

US-10-004-219B-2  
Sequence 2, Application US/10004219B  
Publication No. US20030087414A1  
GENERAL INFORMATION:  
APPLICANT: Macrozyme  
APPLICANT: Aerts, Johannes M.F.G.  
APPLICANT: Boot, Rolf G.  
TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
TITLE OF INVENTION: which mucus is involved or infection diseases  
FILE REFERENCE: 2183-5136US  
CURRENT APPLICATION NUMBER: US/10/004,219B  
CURRENT FILING DATE: 2001-11-02  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2

LENGTH: 1625  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human  
OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid  
OTHER INFORMATION: sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: human AMCase  
OTHER INFORMATION: cDNA sequence and deduced amino acid sequence  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (104) .. (1531)  
US-10-004-219B-2

Alignment Scores:  
Pred. No.: 5.63e-234 Length: 1625  
Score: 2089.50 Matches: 369  
Percent Similarity: 90.09% Conservative: 40  
Best Local Similarity: 81.28% Mismatches: 42  
Query Match: 83.81% Indels: 3  
DB: 9 Gaps: 1



Best Local Similarity: 44.18%  
Query Match: 39.81%  
DB: 9

Mismatches: 137  
Indels: 50  
Gaps: 8

US-10-004-219B-9 (1-452) x US-10-097-340-44 (1-1925)

```
QY 1 TyrAsnLeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
DB 190 TACAAACTGGTCTGCTACTACACCACTGGTCCAGTACCGGAGGCGGATGGAGCTGC 249
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMet 40
DB 250 TTCCAGATGCTCCCTGACCGCTCTCTGTACCATCATCATACAGCTTTGCAATATA 309
QY 41 GlnAsnAsnGluileThrThrileGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
DB 310 AGCAACGATCACATCGACACCTGGGAGTGAATGATGACGCTCTACCGCATGCTCAAC 369
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyGlyTrpAsnPhe 80
DB 370 ACACCTCAAGAACAGAAACCCCACTGAGACTCTCTGTCTCGGAGATGGAATTT 429
QY 81 GlyThrAlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheileThr 100
DB 430 GGCTCTCAAGATTTTCAAGATAGCTTCCAAACACCCAGAGTCCGCGGACTTTTCATCAAG 489
QY 101 SerValileLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 490 TCAGTACCGCATCTCTGCGCACCCAGGCTTTGATGGGCTGGACCTTCCTGGCTCTAC 549
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
DB 550 CTTGGACGGAGA-----GACAAACAGCATTTTACCACCTTAATCAAGGA 594
QY 141 MetArgGluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThr 160
DB 595 ATGAAGCGCGAATTTATAAAGGAAGCC---CAGCCAGGGAAGAGCAGCTCTCTGCTCAGC 651
QY 161 AlaAlaValAlaGlyLysSerAsnileGlnAlaGlyTyrGluileProGluLeuSer 180
DB 652 GCAGCACTGTCTCGGGGAAGTCAACATTCAGCAGCATGATGACATGCCAAGATATCC 711
QY 181 LysTyrLeuAspPheileHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
DB 712 CRAACCTGGATTTTATAGCATCATGACCTACGATTTTCATGAGACCTGCGGTGGACC 771
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
DB 772 ACAGGCCATCACAGTCCCTGTTCCGAGGTGAGGAGTCAAGTCTTCACAGATTCAGC 831
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuile 240
DB 832 AACACTGACTATGCTGGGTACATGTTGAGGCTGGGGCTCTCCAGTAACTGAGCTGGTG 891
QY 241 ValGlyPheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyile 260
DB 892 ATGGGATCCCACTTCGGGAGGACCTTCACTCTGGCTTCT---TCTGAGACTGGTGT 948
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
DB 949 GGAGCCCCCAATCTCAGACCGGGGAATTCAGCGCGGTTCACCAAGGAGGAGCGGACCTT 1008
QY 281 AlaTyrTrpLysLeuThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
DB 1009 GCTACTATGAGATCTGACTTCTCTCCCG---GGAGCCACAGTCCATAGAACCTTCGGC 1065
QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnileLysSer 320
DB 1066 CAGCAGTCCCTATGCCCAAGGCGCAACAGTGGTAGGATACGACGACGAGGAAGC 1125
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetileTrpAla 340
DB 1126 GTCAAAAGCAAGGTGAGTACCTGAGAGGATAGGACGCTGGCAGGCGCCATGATGGGCC 1185
```

```
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGly---LysPheProLeuThr 359
DB 1186 CTGGACCTGGATGACTTCAGGGCTCTCTTCTGGCGCAGGATCTCGGCTTCCCTCTCACC 1245
QY 360 SerThrLeuAsnLysAlaLeuGlyLysSerThrGluGlyCysThrAlaProAspValPro 379
DB 1246 AATGCCATCAAGGATGCATCGCTGC-AACGTAGCCCTCTGTCTGCACACAGCAGCGGG 1304
QY 380 SerGluProValThrProProGlySerGlySerGlyGlyGly----- 394
DB 1305 GCCAAGGATGCCCGCTCCCTCTGCTGCTCCAGTCCGCGGAGCCTGATCACTGCCCTG 1364
QY 394 ----- 394
DB 1365 CTGAGTCCAGCTCAGCTCAGTCTCCCTCTGGGGCCTATGCAGAGTCCACAACA 1424
QY 395 -----SerSerGlyGlySerSerGlySerGlyPheCysAlaAsp--- 408
DB 1425 CACAGATTGAGCTCAGCCCTGGTGGCGAGAGAGTAGGATGGGC-TGTGGGATAGT 1483
QY 409 -----LysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAlaPhe 423
DB 1484 GAGGCATCGCAATGTAAGACTCGGAGATTAGTACACACTTGTGTGATGAT---TAA 1534
QY 424 TrpGlnCysIle 427
DB 1535 TGGAAATGTTTA 1546
```

## RESULT 4

US-10-202-436A-14  
; Sequence 14, Application US/10202436A  
; Publication No. US20030049261A1  
; GENERAL INFORMATION:  
; APPLICANT: ELIAS, JACK A  
; APPLICANT: ZHU, ZHOU  
; TITLE OF INVENTION: METHODS COMPOSITIONS AND KITS RELATING TO CHITINASES AND CHITINASE  
; FILE OF INVENTION: MOLECULES AND INFLAMMATORY DISEASE  
; FILE REFERENCE: 044574-5107  
; CURRENT APPLICATION NUMBER: US/10/202,436A  
; PRIOR FILING DATE: 2002-10-17  
; PRIOR APPLICATION NUMBER: U.S. 60/307,432  
; PRIOR FILING DATE: 2001-07-24  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 1038  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: AMCase sense in situ hybridization probe  
US-10-202-436A-14

Alignment Scores:  
Pred. No.: 1,696-105 Length: 1038  
Score: 987.50 Matches: 183  
Percent Similarity: 80.69% Conservative: 26  
Best Local Similarity: 70.66% Mismatches: 41  
Query Match: 39.61% Indels: 9  
DB: 9 Gaps: 3

US-10-004-219B-9 (1-452) x US-10-202-436A-14 (1-1038)

```
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsn 221
DB 2 GGAGAGAACACCCCTCTACAAATACCCGACTGACCCGCGCAGCAGCTTACCTCAAT 61
QY 222 ValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileVal 241
DB 62 GTGGATTATGTCATGAACACTACTGGAAGGACATGGAGCACCAGCTGAGAGCTCATCGTT 121
QY 242 GlyPheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyileGly 261
```

Db 122 GGATTCCTTACCTATGACACAACTTCTCTGAGCAACCCCTCCACACACTCGGAATGGT 181  
Qy 263 AlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAla 281  
Db 182 GCGCCACCTCTGGTGTCTGCTGCGCCCTATGCCAAGGAGTCTGGGATCTGGGCT 241  
Qy 282 TyrTyrGluLeuCysePheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGln 301  
Db 242 TACTACGAGATCTGTACCTTCTGAAAAATGAGCCACTCAGGGATGGATGCCCTCAG 301  
Qy 302 GluValProTyrAlaTyrIlysalenGluTrpLeuGlyTyrAspAlaIlysalenPhe 321  
Db 302 GAAAGTGCCTTATGATCAGGCAATGTGGTGGCTATGCAACGCTCAAGAGCTTC 361  
Qy 322 SerValIlysalenGluTrpLeuIlysalenGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 341  
Db 362 GATATTAGGCTCAATGGCTTAAGGACACAAATCTGGAGGCGCCATGTCTGGGCCATT 421  
Qy 342 AspLeuAspAspPheThrGlySerPheCysAspGlnGlyPheProLeuThrSerThr 361  
Db 422 GATCTGGATGACTTCACTGGCACTTCTGCAACAGGCAAGTTTCCCTAATCTCCACC 481  
Qy 362 LeuAsnIlysalenGlySerThrGluGlyCysThrAlaProAspValProSerGlu 381  
Db 482 CTGAAGAAGGCGCTCGGCTGAGAGTGAAGTTGCACGCTCCAGCTCAGCCATTGAG 541  
Qy 382 ProValThrThrProPro-----GlySerGlySerGlyGlySerGlyGly 398  
Db 542 CCAATAACTGCTGCTCCAGTGGAGCGGGAACGGAGCGGAGTAGTAGCTCTGGAGGC 601  
Qy 399 SerSerGlyGlySerGlyPhe---CysAlaAspIlysalenGlyLeuTrpProValAla 417  
Db 602 AGCTCGGAGGAGTGGATCTTGTGCTGGCAGCAGCAACAGCTCTAACCCCTGGGC 661  
Qy 418 Asp----AspArgAsnAlaPheTrp-GlnCysIleAsnGlyIleThrTyrGln-GlnHis- 435  
Db 662 AAATTACCAGAGATGCTCTTCTGGGCACTGCTGTAAGTGAAGTCACTACAGGCAAGT 721  
Qy 436 CysGlnAlaGlyLeuVal-PheAspThrSerCysAsnCysCysAsnTrp 451  
Db 722 TGCCAGGCGGGCTGTCTCTCGACACAGCTGTGAATCTGCAACTGG 770

## RESULT 5

US-09-215-077A-4  
; Sequence 4, Application US/09215077A  
; Patent No. US20020031793A1  
; GENERAL INFORMATION:  
; APPLICANT: PRICE, PAUL A.  
; APPLICANT: JOHANSEN, JULIA S.  
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF  
; TITLE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES  
; FILE REFERENCE: 4077-89541US  
; CURRENT APPLICATION NUMBER: US/09/215,077A  
; PRIOR FILING DATE: 1998-12-18  
; PRIOR APPLICATION NUMBER: 08/581,527  
; PRIOR FILING DATE: 1996-04-17  
; PRIOR APPLICATION NUMBER: 08/089,989  
; PRIOR FILING DATE: 1993-07-09  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: cDNA nucleotide  
; OTHER INFORMATION: sequence for the coding region of the gene for  
; OTHER INFORMATION: YKL-40.  
US-09-215-077A-4

Alignment Scores:  
Pred. No.: 1.01e-102 Length: 1681  
Score: 966.50 Matches: 187

Percent Similarity: 66.07% Conservative: 70  
Best Local Similarity: 48.07% Mismatches: 123  
Query Match: 38.77% Indels: 10  
DB: 10 Gaps: 5  
US-10-004-219B-9 (1-452) x US-09-215-077A-4 (1-1681)  
Qy 1 TyrAsnLeuIleCysePheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20  
Db 135 TACAACTGGTCTGCTACTACACAGCTGGTCCAGTACCGGAGGCGGATGGGAGCTGC 194  
Qy 21 LysProAspAlaIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 195 TTCCAGATGCTTACCGCTTCTGTGTATGCCACATCATCATCAGCTTTGCCAATATA 254  
Qy 41 GlnAsnAsnGluIleThrIleGluTrpAsnAspValThrLeuTyrIlysalenPhe 60  
Db 255 AGCAACGATCATCGACACCTGGAGTGGAAATGATGACGCTCTACGGCATGCTCAAC 314  
Qy 61 AspLeuIlysalenAsnArgAsnSerLysLeuIlysalenThrLeuAlaIleGlyGlyTrpAsnPhe 80  
Db 315 ACATCAACAACACGACCCCAACCTGAAGACTCTTGTCTGTCGAGGATGGAACTTT 374  
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100  
Db 375 GGTCTCAAGATTTCACAGATAGCTCCAAACACCCAGAGTCCCGGACTTTTCATCAAG 434  
Qy 101 SerValIleIlysalenPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 435 TCAGTACCGCCATTCTCGCCACCCATGGCTTGTATGGCGGTGACCTTGCCTGGCTCTAC 494  
Qy 121 ProGlySerArgGlySerProProGlnAspIlysalenPheThrValLeuValIlysalen 140  
Db 495 CTGAGCGGAGA-----GACAACAACCATTTTACCACCTAATCAAGGNA 539  
Qy 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160  
Db 540 ATGAAGCGCGAATTATTAAGGAGAGCC--CAGCCAGGAGGAAAAAGCAGCTCTCGTCCAG 596  
Qy 161 AlaAlaValAlaGlyIlysalenSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
Db 597 GCAGCACTGTCTCGCGGGGAGGTCAACATTGACAGCAGCTATGACATTGCCAAGATATCC 656  
Qy 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
Db 657 CAACACCTGGATTCATTAGCATCATGACCTACGATTTTCATGGCGCTCGGCTGGGACC 716  
Qy 201 ThrGlyGluAsnSerProLeuTyrIlysalenTyrProThrGluThrGlySerAsnAlaTyrLeu 220  
Db 717 ACAGGCCATCATCAGTCCCTCAGGCGAGGTTCAGGAGGATGCAAGTCTGACAGATTCCAGC 776  
Qy 221 AsnValAspTyrValMetAsnTyrTrpIlysalenAsnGlyAlaProAlaGluIlysalen 240  
Db 777 AACACTGACTATGCTGTGGGTGATCATGTGAGGCTGGGGGCTCTCCGCAAGTAAAGTGGTG 836  
Qy 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260  
Db 837 ATGGCATCCCACTTCGGAGGAGGACTTCACTCTGGCTTCT--TCTGAGACTGGTGT 893  
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280  
Db 894 CCAGCGCCAACTCTCAGGACCGGAAATCCAGCGCGTTCACCAAGGAGGACGAGGCCCTT 953  
Qy 281 AlaTyrTyrGluIleCysePheThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300  
Db 954 GCCTACTATGATCTGTGACTTCTCTCGC---GGAGCCACACAGTCCATAGAACCTCGGC 1010  
Qy 301 GlnGluValProTyrAlaTyrIlysalenAsnGluTrpLeuGlyTyrAspAsnIlysalen 320  
Db 1011 CAGCAGGTCCTTATGCCACCAAGGCAACAGTGGTAGGATACGACGACGAGGAAAGC 1070  
Qy 321 PheSerValIlysalenGlnTrpLeuIlysalenGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340

Db 1071 GTCAAGCAAGGTGACGTACCTGAAGGATAGGACGTGGCAGGCCCATGGTATGGGCC 1130  
QY 341 ILeAspLeuAspPheThrGlySerPheCysAspGlnGly---LysPheProLeuThr 359  
Db 1131 CTGGACCTGGATGACTTCAGGGCTCTCTGCGGAGGATGCGGCTCTCTCACC 1190  
QY 360 SerThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValPro 379  
Db 1191 AATGCCATCAAGGATGACCTCGCTGC-AAGGTAGCCCTCTGTTCTGCACACACGACGGG 1249  
QY 380 SerGluProValThrThrProProGly 388  
Db 1250 GCAAGGATGCCCGCTCCCGCTGGC 1276

## RESULT 6

US-09-262-213A-4  
; Sequence 4, Application US/09262213A  
; Patent No. US20020090658A1  
; GENERAL INFORMATION:  
; APPLICANT: PRICE, PAUL  
; APPLICANT: JOHANSEN, JULIA  
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF  
; TITLE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES  
; FILE REFERENCE: 407T-895412US  
; CURRENT APPLICATION NUMBER: US/09/262,213A  
; PRIOR FILING DATE: 1999-03-04  
; PRIOR APPLICATION NUMBER: US 08/089,989  
; PRIOR FILING DATE: 1993-07-09  
; PRIOR APPLICATION NUMBER: PCT/US94/07754  
; PRIOR FILING DATE: 1994-07-08  
; PRIOR APPLICATION NUMBER: US 08/581,527  
; PRIOR FILING DATE: 1996-04-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 4  
; LENGTH: 1681  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-262-213A-4

Alignment Scores:  
Pred. No.: 1,01e-102 Length: 1681  
Score: 966.50 Matches: 187  
Percent Similarity: 66.07% Conservative: 70  
Best Local Similarity: 48.07% Mismatches: 123  
Query Match: 38.77% Indels: 10  
DB: 10 Gaps: 5

US-10-004-219B-9 (1-452) x US-09-262-213A-4 (1-1681)

QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20  
Db 135 TACAACCTGGTCTGCTACTACACAGCTGGTCCAGTACCGGAGGCGCATGGAGCTGC 194  
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 195 TTCCACAGATGCCCTTACCGCTCTCTGTGTACCCACATCATCTACAGCTTTCACATATA 254  
QY 41 GlnAsnAsnGluIleThrThrileGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60  
Db 255 AGCAACGATCATCATGACACCTCGGGAGTGGATGATGTGACGCTCTACGGCATGCTCAAC 314  
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPhe 80  
Db 315 ACATCAACAAACAGNACCCACCTGAGATCTCTTGTGTGCGGAGATGGAACTTT 374  
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100  
Db 375 GGCTCTCAAGATTTTCCAAAGATAGCTCCAAACACACAGAGTGGCGGACTTTCATCAAG 434  
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 435 TCAGTACCCGCCCTTCTGCGCACCCCATGGCTTTGATGGGCGTGACCTTGGCTCTAC 494

QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140  
Db 495 CTGGACGGAGA-----GACAAACACCATTTTACCACCTAATCAAGGNA 539  
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnAArgProArgLeuMetValThr 160  
Db 540 ATGAAGGCGCAATTTATAAAGGAGCC---CAGCCAGGAGAAAAGCAGCTCTGCTCAGC 596  
QY 161 AlaAlaValAlaGlyGlyLysSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
Db 597 GCAGCACTGTCTGCGGGGAAGGTCAACATTCAGCAGCATGATGACATTCACCAAGATATCC 656  
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluLysTyr 200  
Db 657 CAACACCTGGATTTTCATTAGCATCATGACCTACGATTTTCATGGCGCTGGCGTGGACC 716  
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220  
Db 717 ACAGGCCATCAGTCCCTCAGGCGAGGTCAAGGATGCAAGTCTGACAGATTCAGC 776  
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlnAlaProAlaGluLysLeuIle 240  
Db 777 AACACTGACTATGCTGGGGTACATGTTGAGGCTGGGGCTCTCCAGTAAGCTGGT 836  
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260  
Db 837 ATGGCATCCCACTTCGGGAGGAGCTTCACTCTGGCTTCT---TCTGAGACTGGTGT 893  
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280  
Db 894 CCAGGCCAATCTCAGGACCGGAAATCCAGGCCGCTTCCACCAAGGAGGAGGACCCCTT 953  
QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300  
Db 954 GCCTACTATGAGATCTGTGACTTCTCTCCGC---GGAGCCACAGTCCATAGAACCTCCGC 1010  
QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGlnTrpLeuGlyTyrAspAsnIleLysSer 320  
Db 1011 CAGCAGGTCCCTATGCCCAAGGCAACCACTGGTGGTAGGATAGCAGCAGCAGGAGG 1070  
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340  
Db 1071 GTCAAAAGCAAGGTGACGTACCTGAAGGATAGGAGTGGCAGGCGGCGCATGTGGGCC 1130  
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGly---LysPheProLeuThr 359  
Db 1131 CTGGACCTGGATGACTTCCAGGCTCTCTTCTGCGGCCAGGATCTGGCTTCTCTCACC 1190  
QY 360 SerThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValPro 379  
Db 1191 AATGCCATCAAGGATGACCTCGCTGC-AACGTAGCCCTCTCTGTCACACACGACGGG 1249  
QY 380 SerGluProValThrThrProProGly 388  
Db 1250 GCAAGGATGCCCGCTCCCGCTGGC 1276

## RESULT 7

US-09-822-830A-402  
; Sequence 402, Application US/09822830A  
; Patent No. US20020142952A1  
; GENERAL INFORMATION:  
; APPLICANT: Genetics Institute, Inc.  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulukota, Kamalakar  
; APPLICANT: Graham, James R.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6402

Db 721 TA

[illegible]

Best Local Similarity: 44.44%

Alignment Scores:			
pred. No.:	1.01e-89	Length:	1474
Score:	854.50	Matches:	180
Percent Similarity:	64.59%	Conservative:	82
Best Local Similarity:	44.44%	Mismatches:	123
Query Match:	34.28%	Indels:	20

[illegible]



Db 220 GAAAAACGTGGTTATGAACACTTTCACAACTTGGGATTTGAAGAATCCAGAATTAACCAACC 279  
Qy 72 LeuLeuAlaIleGlyTyrAsnPhcGlyThrAlaProPheThrMetValSerThr 91  
Db 280 ATGATTTCACTTGGTGGTGTATGAAGCTCGGAAAAATATTCGATATGGCTGCAAT 339  
Qy 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlyPhe 111  
Db 340 CCAACATATCGTCAACAAATTCATCAATCAGTTTGGACTTTTTCGAAGAATACAAGTTC 399  
Qy 112 AspGlyLeuAspLeuAspTyrGlnTyrProGlySerArg---GlySerProGlnAsp 130  
Db 400 GACGGCTCTAGATTGGATTGGAGATCTCGATCTCGATTGGTAAACCCGAAAAATCGAT 459  
Qy 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150  
Db 460 AAACAATACTATTGGCTTTGGTTAGAGACTTAAGAGCGTTTGAACCTCATGCC--- 516  
Qy 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyIleSerAsnIle 170  
Db 517 -----TACTTGTGACTGCTGCAGTATCACCAGGTAAAGACAAATC 558  
Qy 171 GlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190  
Db 559 GACCGAGCTTATGATATCAAGAAATTTGAACAATTTGTTGATTTGGATGAATGTCATGACA 618  
Qy 191 TyrAspLeuHisGlySerTyrGlnTyrGlyGluAsnSerProLeuTyrLysTyr 210  
Db 619 TATGATTACACGGTGGATGGGAAACTTTACGGTCAATGCTCGTTGTATTAACGA 678  
Qy 211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229  
Db 679 CCAGATGAACACTGATGAGTTGCACACTTACTTCAATGTCACTACACCATGCACTATTAT 738  
Qy 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThr 249  
Db 739 TTGAACAATGGTGGCCACGACGACAAATTTGGTAAATGGTGTTCATTTCTATGGCCGTGCT 798  
Qy 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269  
Db 799 TGGACATTTGAAGATCGAAGCAACTCAAACTTGAGATCGACCAAGGCAATGTCGCC 858  
Qy 270 AlaGlyAlaTyrThrArgGlnAlaGlyPheThrAlaTyrTyrGluIleCysThrPheLeu 289  
Db 859 CCAGGTTTCATTTCTGCTGAAGAGTGCTCTCTCATATATAGATTGTGTCAATGTTT 918  
Qy 290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305  
Db 919 CAAAA-----GAAGAATGGCATATCCAAATACGATGAATATTAACAATGCTCCATAT 969  
Qy 306 AlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325  
Db 970 GGTTCAAATGATAAATCTGGTCTGGTACGATGATCTGGCCAGTATATCATGCAAGTTG 1029  
Qy 326 GlnTrpLeuLysGlnAsnAsnPhcGlyGlyAlaMetIleTrpAlaIleAspLeuAspAsp 345  
Db 1030 GCTTTCTGAAAGAAATAGGCGTTTCTGGTGCATGTTTGGTCATTCGAAATCATGAT 1089  
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrIleuAsnLysAla 365  
Db 1090 TTCAAAGGT---CACTGGCGGA-----CGAAAAATCCATTTGTTGAACAAAGTT 1134  
Qy 366 -----LeuGlyIleSerThr 370  
Db 1135 CATATATGATTAAATGGCGATGAAAGAACTCTTTCGAATGCAATTTGGTCCAAAGTACA 1194  
Qy 371 GluGlyCysThrAlaProAspValProSerGluProValThrThrPro 386  
Db 1195 ACGACACCACTCCAAACGACGACACCCACACCCCGACTACAAACGCCA 1242

RESULT 10

US-10-218-743-22/c

; Sequence 22, Application US/10218743

; Publication No. US20030096779A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine A.  
; APPLICANT: Hunter, Shirley Wu  
; APPLICANT: Weber, Eric R.  
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
; FILE REFERENCE: AL-2-C3  
; CURRENT APPLICATION NUMBER: US/10/218,743  
; PRIOR FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: US/09/292,225  
; PRIOR FILING DATE: 1999-04-15  
; PRIOR APPLICATION NUMBER: 60/098,909  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR APPLICATION NUMBER: 60/085,295  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/098,565  
; PRIOR FILING DATE: 1998-04-17  
; PRIOR APPLICATION NUMBER: 09/062,013  
; PRIOR FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 1608  
; TYPE: DNA  
; ORGANISM: Dermatophagoides farinae  
; US-10-218-743-22  
Alignment Scores:  
Pred. No.: 4,69e-75 Length: 1608  
Score: 729.50 Matches: 154  
Percent Similarity: 53.61% Conservative: 69  
Best Local Similarity: 37.02% Mismatches: 144  
Query Match: 29.26% Indels: 49  
DB: 9 Gaps: 11

US-10-004-219B-9 (1-452) x US-10-218-743-22 (1-1608)

Qy 3 LeuIleCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22  
Db 1563 ATGTTTGTATGTGGACATGGTCCGTATATCATATA---GTTGATCCATACACTATC 1507  
Qy 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42  
Db 1506 GAAGATATTGATCCATTCAAGTGTACACATTAAATGATGTTTCGCTAAAAATTCATGAA 1447  
Qy 43 AsnGluIleThrThr-----IleGluTrpAsn 51  
Db 1446 TACAATAACACAATTCAAAGTTTTCGATCCTTACCAGATGATAACCAATCACTATGG--- 1390  
Qy 52 AspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71  
Db 1389 GAAAAACGTGGTTATGAACGTTTCAACAACTTGGATTTGAAGAATCCAGAAATTAACCAACC 1330  
Qy 72 LeuLeuAlaIleGlyTyrAsnPhcGlyThrAlaProPheThrThrMetValSerThr 91  
Db 1329 ATGATTTCACTTGGTGGTGTATGAAGCTCGGAAAAATATTCGATATGGCTGCAAT 1270  
Qy 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlyPhe 111  
Db 1269 CCAACATATCGTCAACAAATTCATCAATCAGTTTGGACTTTTTCGAAGAATACAAGTTC 1210  
Qy 112 AspGlyLeuAspLeuAspTyrGlnTyrProGlySerArg---GlySerProGlnAsp 130  
Db 1209 GACGGCTCTAGATTGGATTGGAGTATCTCGATCTCGATTGGGTAAACCCGAAAAATCGAT 1150  
Qy 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150  
Db 1149 AAACAATACTATTGGCTTTGGTTAGAGAACTTAAGAGCGTTTGAACCTCATGCC--- 1093  
Qy 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyIleSerAsnIle 170  
Db 1092 -----TACTTGTGACTGCTGCAGTATCACCAGGTAAAGACAAATC 1051

```
QY 171 GlnAlaGlyTyrGluLeuProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
Db 1050 GACCGAGCTTATGATCAAGAAATGAACAATTTCTCGATGGATGATGTCATGACA 991
QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGluAlaAsnSerProLeuTyrLysTyr 210
Db 990 TATGATTACACGGTGGATGGGAAACTTTTACGGGTACAAATGCTCCGTTGTATAAACGA 931
QY 211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
Db 930 CCAGATGAACATGATGAGTGGACATCTTCAATGTCAATGTACACCATGACATATTAT 871
QY 230 LysAsnAsnGlyAlaProAlaGluLysLeuLeuValGlyPheProGluTyrGlyHisThr 249
Db 870 TTGAACATGGTGGCCACGACAGACAATTTGTAAGGTGTCCATCTATGGCGGTCT 811
QY 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
Db 810 TGGAGCATTGAAGATCGAAGCAAACTCAAACTTGGAGATCCAGCCAAAGCATGTCGCC 751
QY 270 AlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluIleCysThrPheLeu 289
Db 750 CCAGGTTTCATTCTCGTGAAGAAGGTGCTCTCATATATAGAAATGTGTCAATGTGT 691
QY 290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305
Db 690 CAAAAA-----GAAGANTGGCATATCCAAATGATGATGATGATGATGATGATGAT 640
QY 306 AlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
Db 639 GGTTCATGATGAATAAATCTGGTGGTTCAGATGATGATGATGATGATGATGATGATGAT 580
QY 326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAsp 345
Db 579 GCTTTCTCTCAAGAAATAGCGGTTCTGTGTGTCATGTTTGGTTCATGGAATAATGATGAT 520
QY 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
Db 519 TTCAAAGGT---CACTGCGGA-----CCGAAAAATCCATTGTTGAACAAAGTT 475
QY 366 -----LeuGlyIleSerThr 370
Db 474 CATATATGATTAATGGCGATGAAAGAACTCTTTGCAATGCAATTTTGGTTCGCAAGTACA 415
QY 371 GluGlyCysThrAlaProAspValProSerGluProValThrThrPro 386
Db 414 ACGACACCAACTCCACGACGACACCCACACCCCGACTACACGCGCA 367

RESULT 11
US-10-218-743-17
; Sequence 17, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1665)
US-10-218-743-17

Alignment Scores:
Score: 4.95e-75 Length: 1665
Pred. No.: 729.50 Matches: 154
Percent Similarity: 53.61% Conservative: 69
Best Local Similarity: 37.02% Mismatches: 144
Query Match: 29.26% Indels: 49
DB: 9 Gaps: 11

US-10-004-219B-9 (1-452) x US-10-218-743-17 (1-1665)
QY 3 LeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
Db 103 ATGTTTGTATGTTGGAAATGTCCTGATATATCAAAA---GTTGATCCATACACTATC 159
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
Db 160 GAAGATATTGATCCATTCAAGTGTACACATTAAATGATGTTTCGCTTAAATTTGATGAA 219
QY 43 AsnGluIleThrThr-----IleGluTrpAsn 51
Db 220 TACAATACACAAATTCAGTTTTCGATCTTACCAGATGATAACCATACATCATGG--- 276
QY 52 AspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71
Db 277 GAAAAACGTGGTTATGAACGTTTCAACACTTGCATGGAAGATCCAGAAATTAACACC 336
QY 72 LeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPheThrMetValSerThr 91
Db 337 ATGATTTCACCTGGTGGTTGGTATGAAGCTCGGAAAAATATATCCGATATGCTGCAAA 396
QY 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlyPhe 111
Db 397 CCAACATATCGTCAACAATTCATCAATCAGTTTGGACTTTTTCGAAGAATACAAAGTTC 456
QY 112 AspGlyLeuAspLeuAspTrpGluTyrProGlySerArg---GlySerProProGlnAsp 130
Db 457 GACGGTCTAGATTGGATTGGAGTATCTCGATTCGATGGGTAAACCGGAAATTCAT 516
QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150
Db 517 AAACAAACTATTTGGCTTTGGTTAGAGAACTTAAAGACGCTTTTGAACCTCATGGC--- 573
QY 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyLysIleSerAsnIle 170
Db 574 -----TACTTGTGACTGCTGAGTATCACCAGGTAAACAGCAAAATC 615
QY 171 GlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
Db 616 GACCGAGCTTATGATCAAGAAATGAACAATTTGTCGATGGATGAATGTCATGACA 675
QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLysTyr 210
Db 676 TATGATTACACCGTGGATGGGAAACTTTTACGGGTACAAATGCTCCGTTGTATAAACGA 735
QY 211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
Db 736 CCAGATGAACATGATGAGTTCACACTTACTTCAATGTCACCATCACCATGACATATTAT 795
QY 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThr 249
Db 796 TTGAACATGGTGGCCACGACAGACAATTTGTAAGGTGTCCATTTCTATGGCGGTGCT 855
QY 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
```

Db 856 TGGAGCATTTGAAGATCGAAGCAAACTCAAACCTGGAGATCCAGCCAAAGCATGCGCCC 915  
Qy 270 AlaGlyAlaTyThrArgGlnAlaGlyPheTrpAlaTyThrGluLeuCysThrPheLeu 289  
Db 916 CCAGGTTTCATTCTGGTGGGAAGGTGCTCTCATATATAGAAATTTGTCGAATGTTT 975  
Qy 290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305  
Db 976 CAAAAA-----GAAGATGGCATATCCAAATACGATGAATATTACAAATCCTCCATAT 1026  
Qy 306 AlaTyLysAlaAsnGluTrpLeuGlyTyThrAspAsnIleLysSerPheSerValLysAla 325  
Db 1027 GGTTCACATGATATAAATCTGGTGGTACGATGATCTGCCAGTATATCATGCAAGTTG 1086  
Qy 326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspAsp 345  
Db 1087 GCTTTCCTGAAAGAAATAGGCGTTCTGTCATGCTGTTGTCATGGAATAATGATGAT 1146  
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365  
Db 1147 TTCAAAGGT---CACTGCGGA-----CCGAAAAATCCATTGTTGAACAAAGTT 1191  
Qy 366 -----LeuGlyLysSerThr 370  
Db 1192 CATAATATGATTAAATGGCGATGAAAGAACTCTTTTGAATGCAATTTTGGGTCCAAAGTACA 1251  
Qy 371 GluGlyCysThrAlaProAspValProSerGluProValThrPro 386  
Db 1252 ACGACACCACTCCAAACGACGACCAACACCCCGCACTACAAACGCCA 1299

## RESULT 12

US-10-218-743-19/c  
; Sequence 19, Application US/10218743  
; Publication No. US20030096779A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine A.  
; APPLICANT: Hunter, Shirley Wu  
; APPLICANT: Weber, Eric R.  
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: AL-2-C3  
; CURRENT APPLICATION NUMBER: US/10/218,743  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: US/09/292,225  
; PRIOR FILING DATE: 1999-04-15  
; PRIOR APPLICATION NUMBER: 60/098,909  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR APPLICATION NUMBER: 60/085,295  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/098,565  
; PRIOR FILING DATE: 1998-04-17  
; PRIOR APPLICATION NUMBER: 09/062,013  
; PRIOR FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 1665  
; TYPE: DNA  
; ORGANISM: Dermatophagoides farinae  
US-10-218-743-19

## Alignment Scores:

Pred. No.: 4,95e-75 Length: 1665  
Score: 729.50 Matches: 154  
Percent Similarity: 53.61% Conservative: 69  
Best Local Similarity: 37.02% Mismatches: 144  
Query Match: 29.26% Indels: 49  
DB: 9 Gaps: 11

US-10-004-219b-9 (1-452) x US-10-218-743-19 (1-1665)

Qy 3 LeuIleCysTyThrPheThrAsnTrpAlaGlnTyThrArgProGlyLeuGlySerPheLysPro 22

Db 1563 ATTGTTTGTATTGTTGGAACATGGTCCGTATATCATATAA---GTTGATCCATACACTATC 1507  
Qy 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyThrAlaPheAlaGlyMetGlnAsn 42  
Db 1506 GAAGATATTGATCCATTCAGTGTACACATTTAATGTATGTTTTCGCTAAAATGATGAA 1447  
Qy 43 AsnGluIleThrThr-----IleGluTrpAsn 51  
Db 1446 TACAAATACACAATTCAAAGTTTTCGATCCTTACCAAGATGATAACCAATCACTATG--- 1390  
Qy 52 AspValThrLeuTyLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71  
Db 1389 GAAAAACGCTGGTTATGAACGTTTCAACAACTTCGCAATGGAAGAATCCAGAATTAACCAACC 1330  
Qy 72 LeuLeuAlaIleGlyGlyTrpAsnPheGlyThrAlaProPheThrThrMetValSerThr 91  
Db 1339 ATGATTTCCATTTGGTGGTTGTATGAAGCTCGGAAAAATATTCGATATGCTGCAAAAT 1270  
Qy 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyThrGlyPhe 111  
Db 1269 CCAACATATCGTCAACAATTCATACAATCAGTTTGGACTTTTTCGAAGAATACAAAGTTT 1210  
Qy 112 AspGlyLeuAspLeuAspTrpGluTyThrProGlySerArg---GlySerProProGlnAsp 130  
Db 1209 GACGGTCTAGATTTGGATTGGAGTATCTCGATCTCGATTGGGTAAACCCGAAAAATCGAT 1150  
Qy 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150  
Db 1149 AAACAAAACATAATTTGGCTTTGGTTAGAGAACTTAAAGACGCTTTTGAACCTCATGGC--- 1093  
Qy 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsnIle 170  
Db 1092 -----TACTTGTGACTGCTGCAGTATCACCAGGTAAAGACAAAATC 1051  
Qy 171 GlnAlaGlyTyThrGluIleProGluLeuSerLysTyThrLeuAspPheIleHisValMetThr 190  
Db 1050 GACCGAGCTTATGATATCAAGAATGAACAAATTTGTCGATTGGATGAATGTCATGACA 991  
Qy 191 TyrAspLeuHisGlySerTrpGluGlyTyThrGlyGluAsnSerProLeuTyThrLysTyThr 210  
Db 990 TATGATTACCAACGGTGGATGGGAAACCTTTACGCTCAATGCTCCGTTGTATAAACGA 931  
Qy 211 ProThrGluThrGlySer---AsnAlaTyThrLeuAsnValAspTyThrValMetAsnTyThr 229  
Db 930 CCAGATGAACATGATGAGTGCACACTTACTTCAATGTCACTACACCATGCACATATTAT 871  
Qy 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyThrGlyHisThr 249  
Db 870 TTGAACAATGGTGGCCACAGAGACAAATTTGGTAAATGGGTGTTCCATTCTATGGCGTGCT 811  
Qy 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269  
Db 810 TGGAGCATTTGAAGATCGAAGCAAACTCAAACTTGGAGATCCAGCCAAAGGCGATGCGCCC 751  
Qy 270 AlaGlyAlaTyThrArgGlnAlaGlyPheTrpAlaTyThrGluIleCysThrPheLeu 289  
Db 750 CCAGGTTTCATTCTGGTGAAGAGTGTCTCTCATATATAGAAATTTGTCATATTGTTT 691  
Qy 290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305  
Db 690 CAAAAA-----GAAGATGGCATATCCAAATACGATGAATATTACAAATCCTCCATAT 640  
Qy 306 AlaTyLysAlaAsnGluTrpLeuGlyTyThrAspAsnIleLysSerPheSerValLysAla 325  
Db 639 GGTTCACATGATATAAATCTGGTGGTACGATGATCTGCCAGTATATCATGCAAGTTG 580  
Qy 326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspAsp 345  
Db 579 GCTTTCCTGAAAGAAATTAGCGGTTTCTGGTGCATGGTTTGGTCATTGGAAAAATGATGAT 520  
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365

Db 519 TTCAAAGGT---CACTGCGGA-----CCGAAAAATCATTGTTGAACAAAGTT 475  
 QY 366 -----LeuGlyIleSerThr 370  
 Db 474 CATAATATGATTAAATGGCGATGAAGAAGAACTCTTTTGAATGCAATTTGGTTCAGATGACA 415  
 QY 371 GluGlyCysThrAlaProAspValProSerGluProValThrPro 386  
 Db 414 ACGACACCAACTCCACGACGACGACCCACACCCCGACTACAGGCCA 367

RESULT 13  
 US-10-218-743-14  
 ; Sequence 14, Application US/10218743  
 ; Publication No. US20030096779A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McCall, Catherine A.  
 ; APPLICANT: Hunter, Shirley Wu  
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
 ; AND USES THEREOF  
 ; FILE REFERENCE: AL-2-C3  
 ; CURRENT APPLICATION NUMBER: US/10/218,743  
 ; CURRENT FILING DATE: 2002-08-13  
 ; PRIOR APPLICATION NUMBER: US/09/292,225  
 ; PRIOR FILING DATE: 1999-04-15  
 ; PRIOR APPLICATION NUMBER: 60/098,909  
 ; PRIOR FILING DATE: 1998-09-02  
 ; PRIOR APPLICATION NUMBER: 60/085,295  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/098,565  
 ; PRIOR FILING DATE: 1998-04-17  
 ; PRIOR APPLICATION NUMBER: 09/062,013  
 ; PRIOR FILING DATE: 1998-04-17  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 14  
 ; LENGTH: 1752  
 ; TYPE: DNA  
 ; ORGANISM: Dermatophagoides farinae  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(1665)  
 ; US-10-218-743-14

Alignment Scores:  
 Pred. No.: 5,35e-75 Length: 1752  
 Score: 729.50 Matches: 154  
 Percent Similarity: 53.61% Conservative: 69  
 Best Local Similarity: 37.02% Mismatches: 144  
 Query Match: 29.26% Indels: 49  
 DB: 9 Gaps: 11

US-10-004-219B-9 (1-452) x US-10-218-743-14 (1-1752)

QY 3 LeuIleCysThrPheThrAsnTrpAlaGlnTrpArgProGlyLeuGlySerPheIysPro 22  
 Db 103 ATTGTTGTTGTTGTTGGAACATGTTCCGATATATCAAAA---GTTGATCCATACACTATC 159  
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42  
 Db 160 GAAGATATTGATCCATTCAAGTGTACACATTTAATGTTATGTTGTTTCCGTAATAATGATGAA 219  
 QY 43 AsnGluIleThrThr-----IleGluTrpAsn 51  
 Db 220 TACAATAACACAAATCAAGTTTTCGATCTCTTCCAGATGATAACCATCACTATG--- 276  
 QY 52 AspValThrLeuTyrIlyAlaPheAsnAspLeuIleAsnArgAsnSerIlyLeuIlyThr 71  
 Db 277 GAAAAACGGTGTATGAACGTTTCAACACTTCCGATTCAGAAATCCAGAAATTAACACC 336  
 QY 72 LeuLeuAlaIleGlyTyrTrpAsnPheGlyThrAlaProPheThrThrMetValSerThr 91  
 Db 337 ATGATTTTCACTGGTGGTGGTATGAAGGCTCGGAAAAATATTTCCGATATGCTGCAAT 396

QY 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTrpGlyPhe 111  
 Db 397 CCAACATATCGTCAACATTTTATACAAATCAGTTTGGACTTTTTCAGAAATACAAAGTTC 456  
 QY 112 AspGlyLeuAspLeuAspTrpGluTyrProGlySerArg---GlySerProProGlnAsp 130  
 Db 457 GACGGTCTAGATTGGATTGGAGTATCTCGATCTCGATTGGGTAAACCCGAAATATCATGAT 516  
 QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150  
 Db 517 AAACAAACTATTGGCTTTGGTTAGAGAACTTAAAGACGCTTTTGAACCTCATGGC--- 573  
 QY 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsnIle 170  
 Db 574 -----TACTTGTGACTGCTGCAGTATCACCAGGTAAAGACAAATC 615  
 QY 171 GlnAlaGlyTyrGluIleProGluLeuSerIlySerIlyLeuAspPheIleHisValMetThr 190  
 Db 616 GACCGAGCTTATGATATCAAGAATTTGAACAAATTTGTCGATTGGATGAATGTCATGACA 675  
 QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrIlySer 210  
 Db 676 TATGATTACCACGGTGGATGGGAAACTTTTACGGTCACAATGCTCCGTTGTATAAACGA 735  
 QY 211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229  
 Db 736 CCAGATGAACATGATGAGTTCGCACACTTACTTCAATGTCACACTACACCATGACATTTAT 795  
 QY 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThr 249  
 Db 796 TTGAACAATGGTGCACCAGACAAATTTGTAATGGGTGTTCCATTTCTATGGCCGTGCT 855  
 QY 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269  
 Db 856 TGGAGCATTTGAAGATCGAAGCAAACTCAAACTTGGAGATCCAGCCAAAGGCGATGCGCCC 915  
 QY 270 AlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluIleCysThrPheLeu 289  
 Db 916 CCAGGTTTCAATTTCTGGTGAAGAGGTGCTCTCATATATAGAAATTTGTCATTTGTT 975  
 QY 290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305  
 Db 976 CAAAAA-----GAAGAATGGCATATCCAATACGATGAATATTACAATGCTCCATAT 1026  
 QY 306 AlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325  
 Db 1027 GGTTCAAATGATAAAATCTGGGTGCGTTTACCATGATCTGGCCAGTATATCATGCAAGTTG 1086  
 QY 326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspAsp 345  
 Db 1087 GCTTTCTCTGAAGAATTAGGCGTTTCTGCTGTCATGTTGTTGTCATTTGCGAAATGATCAT 1146  
 QY 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365  
 Db 1147 TTCAAAGGT---CACTGCGGA-----CCGAAAAATTCATTTGCGTCCAAAGT 1191  
 QY 366 -----LeuGlyIleSerThr 370  
 Db 1192 CATATATGATTATGCGGATGAAGAAAGAACTCTTTTTCGATGCAATTTTGGTCCAAAGTACA 1251  
 QY 371 GluGlyCysThrAlaProAspValProSerGluProValThrPro 386  
 Db 1252 ACGACACCAACTCCACGACGACGACCCACCAACCCCGACTACACGCCA 1299

RESULT 14  
 US-10-218-743-16/c  
 ; Sequence 16, Application US/10218743  
 ; Publication No. US20030096779A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McCall, Catherine A.  
 ; APPLICANT: Hunter, Shirley Wu  
 ; APPLICANT: Weber, Eric R.



FEATURE:  
NAME/KEY: CDS  
LOCATION: (1) ... (1470)  
US-10-218-743-40

Alignment Scores:  
Pred. No.: 3,07e-74 Length: 1470  
Score: 722.00 Matches: 158  
Percent Similarity: 53.96% Conservative: 67  
Best Local Similarity: 37.89% Mismatches: 142  
Query Match: 28.96% Indels: 50  
DB: 9 Gaps: 13

US-10-004-219B-9 (1-452) x US-10-218-743-40 (1-1470)

```
QY 3 LeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 46 ATCGTATGTTATGTTGGAAACATGGTCGGTTTATCATATAA---GTTGATCCATACACAATT 102
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
DB 103 GAAGATATTGATCCCTTCAAGTACTCAATTGATGATGTTTGGTAAATCGATGAA 162
QY 43 AsnGluIleThr-----leGluTrpAsn 51
DB 163 TACAAATACACCATTCACGTTTTTGATCCATTTCAAGATGATAACCATAACTCATGG--- 219
QY 52 AspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71
DB 220 GAAACACCGGTATGAACGTTTCAACACTTGAGATTGAAGATTCGAAGATTCGACCAACC 279
QY 72 LeuLeuAlaIleGlyLysTrpAsnPheGlyThrAlaProPheThrMetValSerThr 91
DB 280 ATGATTTCAATGGTGGTGGTATGATGAGTTTCAAGAAATATTCGATATGCGACCAAT 339
QY 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlyPhe 111
DB 340 CCAACATATCGTCAGCAATTTGTTCAATCAGTTTGGACTTTTTCGACGATTCGAAATACAAATTC 399
QY 112 AspGlyLeuAsnLeuAspTrpGluTyrProGlySerArg---GlySerProGluLys 130
DB 400 GATGGCTAGATTGGATTGGGAATATCTGGATCAGGTTAGGCAATCTTAAATTCGAT 459
QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150
DB 460 AAACAAACTATTAACTATTAGTAGAAGCTTAAAGAGGCATTTGAACCTTCGCG--- 516
QY 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyLysSerAsnIle 170
DB 517 -----TACTTTGTTGACTGCGGAGTATCACCGGTAAAGATAAAAT 558
QY 171 GlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
DB 559 GACGTAGCTTATGAGCTCAAGAAATTTGAACCAATTTGTTGATGGATGAATGTCTGACT 618
QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyLysAsnSerProLeuTyrLysTyr 210
DB 619 TATGATTACCATGGCGGATGGGAATGTTTCGGCCATAATGCTTCGTTGTATAACGA 678
QY 211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
DB 679 CCGGATGAAACCGATGAATTCACACTTACTTCAATGTCACTACACCATGACATATTAT 738
QY 230 LysAsnAsnGlyValaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThr 249
DB 739 TTGAACAATGGCGCTACTCGAGACAAACTTGTATGGGTGTTCCATTCATGGTCGTGCT 798
QY 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
DB 799 TGGAGCATCGAAGATCGAAGCAAGTCAAACTTGGCGATCCGCCCAAGGCATGCTCTCT 858
QY 270 AlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluIleCysThrPheLeu 289
```

```
DB 859 CCTGGTTTTATTACTGGTGAAGAAGGTGTTCTCTCATATCATCGAATTGTGTCAGTTATTC 918
QY 290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305
DB 919 CAGAAA-----GAAGAATGGCATATTCAATACGATGAATATTACAATGCTCCATAC 969
QY 306 AlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
DB 970 GGATATAATGATAAAATCTGGGTTGGTTACGATGATCTGGCTAGTATATCATGCAAGTTG 1029
QY 326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspAsp 345
DB 1030 GCCTTTCTCAAGAATTTGGCGCTCTCTGGCGTTATGATATGTCATTCGAAACCATGAT 1089
QY 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
DB 1090 TTCAAAGGT---CATTCGGACCG---AAATATCCATTG-----TTGAACAAAGTT 1134
QY 366 -----LeuGlyIleSerThr 370
DB 1135 CACAATATGATCAATGGTGTGATGAAAAAGAACTCTTACGAATGTTCTTTGGGCCCAAGTACA 1194
QY 371 GluGlyCysThrAlaProAspValProSerGluProValThrThrProPro 387
DB 1195 ACCACACCAACA---CCAACACCCCGTCAACTCTTCGACTACCAACCA 1242
```

Search completed: June 30, 2003, 03:57:52  
Job time : 203.82 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:56:54 ; Search time 14.0414 Seconds \*  
(without alignments)  
3094.613 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493  
Sequence: 1 YNLICYFTNWAQYRPLGLSP.....QQHCQAGLVFTSCNCNWP 452

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1334.5	53.5	399	2 S27879	secretory protein
2	1013	40.6	654	2 I38605	oviductal glycopro
3	1010	40.5	537	2 S57197	oviduct-specific g
4	1003	40.2	539	2 I46470	estrogen dependent
5	987.5	39.6	383	2 A49562	cartilage glycopro
6	973.5	39.0	383	2 S51327	heparin-binding gl
7	901	36.1	483	2 A53918	chitinase (EC 3.2.
8	889.5	35.7	617	2 T15408	hypothetical prote
9	863.5	34.6	554	2 A56596	chitinase (EC 3.2.
10	844.5	33.9	525	2 T44445	chitinase (EC 3.2.
11	818.5	32.8	405	2 S61551	breast-regressing
12	810	32.5	504	2 A38221	chitinase (EC 3.2.
13	781.5	31.3	1635	2 T14075	chitinase (EC 3.2.
14	495	19.9	1484	2 T22275	hypothetical prote
15	484	19.4	699	2 A38368	chitinase (EC 3.2.
16	482.5	19.4	1215	2 T43916	chitinase A (impor
17	477.5	19.2	756	2 AB1452	chitinase B homolo
18	475	19.1	599	2 DB3764	chitinase BH0916 (
19	472.5	19.0	756	2 AB1088	chitinase B homolo
20	472.5	19.0	831	2 T00323	chitinase (EC 3.2.
21	470.5	18.9	423	2 TQ1975	chitinase (EC 3.2.
22	470.5	18.9	2025	2 TQ3884	hypothetical prote
23	468.5	18.6	398	2 T04761	chitinase homolog
24	464	18.6	379	2 T04762	chitinase homolog
25	455.5	18.3	424	2 S47133	chitinase (EC 3.2.
26	453.5	18.2	427	2 JC4565	chitinase (EC 3.2.
27	450.5	18.1	424	2 S68121	chitinase I precur
28	442.5	17.7	423	2 S51369	chitinase - fungus
29	418	16.8	378	2 S51591	chitinase (EC 3.2.

30	417.5	16.7	546	2 F84238	chitinase (impor
31	413	16.6	633	2 T24898	hypothetical prote
32	395	15.8	366	2 T04763	chitinase homolog
33	392	15.7	511	2 S61166	probable membrane
34	389	15.6	452	2 JC4038	47K glycoprotein p
35	374	15.0	1054	2 T30933	chitinase (EC 3.2.
36	368	14.8	371	2 T04756	chitinase homolog
37	368	14.8	1051	2 D82428	chitodextrinase VC
38	365.5	14.7	499	2 S04856	chitinase (EC 3.2.
39	364.5	14.6	499	2 S52422	chitinase (EC 3.2.
40	360	14.4	365	2 T04757	chitinase homolog
41	356.5	14.3	332	2 T04754	hypothetical prote
42	355.5	14.3	563	2 S60651	chitinase precursor
43	341.5	13.7	421	2 T04753	hypothetical prote
44	341.5	13.7	1046	2 T30199	chitinase (EC 3.2.
45	333.5	13.4	561	2 A25090	chitinase (EC 3.2.

ALIGNMENTS

RESULT 1

S27879

secretory protein YM-1 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 22-Jun-1999

C:Accession: S27879

R:Chang, N.C.A.; Liu, C.H.; Chang, A.C.

submitted to the EMBL Data Library, June 1992

A:Description: Molecular characterization of a secretory protein (YM-1) transiently expre

A:Reference number: S27879

A:Accession: S27879

A:Molecule type: mRNA

A:Residues: 1-399 <CHA>

A:Cross-references: EMBL:M94584; NID:G202441; PIDN:AB62394.1; PID:G202442

C:Superfamily: Streptomyces chitinase chi40

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-399/Product: secretory protein YM-1 #status predicted <MAT>

Query Match 53.5%; Score 1334.5; DB 2; Length 399;  
Best Local Similarity 64.5%; Pred. No. 7.6e-92; Mismatches 80; Indels 1; Gaps 1;  
Matches 243; Conservative 53;

Qy	1	YNLCYFTNWAQYRPLGLSPKDDINPCLCTHLIYAFAGMONEITTIEMNDVTLYKAFN	60
Db	22	YQLMCYTTSWAKDRPIEGSPKGNIDPCLCTHLIYAFAGMONEITYTHEQDLRDYEALN	81
Qy	61	DLKN-RNSKLTLLAIGQWNGFTAPFTTMVTSQNRQFTITSVKFLQYQPDGLDLWE	119
Db	82	GLKDKKNTLKTLLAIGQWNGFTAPFTTMVTSQNRQFTITSVKFLQYQPDGLDLWE	141
Qy	120	YPSRGSPQDKHLFTVLKEMREAFQEAIESNRPRLMTAAVAGGISNTOAGYEIPBL	179
Db	142	YPSRGSPQDKHLFTVLKEMREAFQEAIESNRPRLMTAAVAGGISNTOAGYEIPBL	201
Qy	180	SKYLDIFHWMTYDLHSGWEGYTGNSPLYKYPTETGNSNAYLNVDYVMYKNGKAPAEKL	239
Db	202	LSLLDYIQWYDLHDPKDYTGNSPLYKSPYDYGKADLNVDLSIISWKDHGAASEKL	261
Qy	240	IVGPEYGHFTILRNPSONGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVDA	299
Db	262	IVGPEYGHFTILRNPSONGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVDA	321
Qy	300	SOEVPYAYKANEWLYGDNKFSVKAQWLKQNNFGAMWAIDLDLDFGSCDQCKEPLT	359
Db	322	FOEVPYAYQGNNEWGYDNVRSFKLKAQWLKQNNLGGAVVWPLDMDDFSGSFCHQRHPLT	381
Qy	360	STLNKALGISTEGCTAP	376
Db	382	STLKGDLNIHSASCKGP	398

RESULT 2

138605

oviductal glycoprotein - human

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 29-May-1998

C:Accession: 138605

R:Arrias, E.B.; Verhage, H.G.; Jaffe, R.C.

Biol. Reprod. 51, 685-694, 1994

A:Title: Complementary deoxyribonucleic acid cloning and molecular characterization of a

A:Reference number: 138605; MUID:95119256; PMID:7819450

A:Accession: 138605

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-654 &lt;RES&gt;

A:Cross-references: EMBL:U09550; NID:9529147; PID:g529148

Query Match 40.6%; Score 1013; DB 2; Length 654;

Best Local Similarity 48.8%; Pred. No. 1.4e-67;

Matches 203; Conservative 58; Mismatches 125; Indels 30; Gaps 5;

QY 1 YNLICVFTWQAQYRPGLSGFKPDDINPCLCTHLIYAFAGMNNETITIEWND-VTLTKAF 59

DB 22 HKLVCTFTWAHSRPGASILPHDLDFLCTHLIFAFASNNNNQIVAKOLQDEKILYPEF 81

QY 60 NDLKRNKSLKTLIAIGGNFCTAPFTTVMVSTSONRQTFTITSVIFKFLQYFGDGLDWE 119

DB 82 NKLKERNRELKTLISIGGNFCTAPFTTVMVSTSONRQTFTITSVIFKFLQYFGDGLDWE 141

QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGGINSIQAGYEIPEL 179

DB 142 YPGLRGSPMDHRTWTFLEILLQAFKNEAQLTWPRLLLSAAVSGDPHVQKADARLL 201

QY 180 SKYLDPIHMTYDLHGSWEVGTGENSPLYKYPTETGSGNAYLNVDYVMYKKNNGAPAEKL 239

DB 202 GRLLDFISVLSDHGSWEKVTGHNPSLFLSGDPKSSA-----YAMNYWROLGVPEPEKL 256

QY 240 IVGFPEYGHFTILRNPDSNGICAPTSGDGPAGAYTRQAGFWAYEICTFLRSGATEVMDA 299

DB 257 IMGIPYGRTHLLKASKNGLOARAGPASPQKTKQAGFLAYFEICFSV-WGAKKHWD 315

QY 300 SEVPYAYKANWGLVDYDKISFVKAQWLKQNNFGAMIAIDDDFTGSCDQKGFPLT 359

DB 316 QYVVPYANKKEWGVGDYDAISFGYKAFKREHFGGAMVWTLDDDFRGYFCGCTGPFPLV 375

QY 360 STLNKALGISTEGCTAPDV-----PSEPVT-----PBG 388

DB 376 VYNDLILVRAEFSSTSLPQFWLSSAVNSSDTPERLAVTIAWTTDSKILPPGEG 431

RESULT 3

S57197

oviduct-specific glycoprotein 95K precursor - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 05-Nov-1999

C:Accession: S57197

R:Sendai, Y.; Abe, H.; Kikuchi, M.; Satoh, T.; Hoshi, H.

Biol. Reprod. 50, 927-934, 1994

A:Title: Purification and molecular cloning of bovine oviduct-specific glycoprotein.

A:Reference number: S57197; MUID:94257768; PMID:8199272

A:Accession: S57197

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-537 &lt;SEN&gt;

A:Cross-references: EMBL:D16639; NID:g391621; PID:BAA04065.1; PID:d1004583; PID:g391622

C:Keywords: glycoprotein

Query Match 40.5%; Score 1010; DB 2; Length 537;

Best Local Similarity 49.2%; Pred. No. 1.8e-67;

Matches 203; Conservative 53; Mismatches 125; Indels 32; Gaps 6;

QY 1 YNLICVFTWQAQYRPGLSGFKPDDINPCLCTHLIYAFAGMNNETITIEWND-VTLTKAF 59

DB 19 HKLVCTFTWAHSRPGASILPHDLDFLCTHLIFAFASNNNNQIVPKDQDEKILYPEF 78

QY 60 NDLKRNKSLKTLIAIGGNFCTAPFTTVMVSTSONRQTFTITSVIFKFLQYFGDGLDWE 119

DB 79 NKLKERNRELKTLISIGGNFCTAPFTTVMVSTSONRQTFTITSVIFKFLQYFGDGLDWE 138

QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGGINSIQAGYEIPEL 179

DB 139 YPGLRGSPMDHRTWTFLEILLQAFKNEAQLTWPRLLLSAAVSGDPHVQKADARLL 198

QY 180 SKYLDPIHMTYDLHGSWEVGTGENSPLYKYPTETGSGNAYLNVDYVMYKKNNGAPAEKL 239

DB 199 GRLLDFISVLSDHGSWEKVTGHNPSLFLSGDPKSSA-----YAMNYWROLGVPEPEKL 253

QY 240 IVGFPEYGHFTILRNPDSNGICAPTSGDGPAGAYTRQAGFWAYEICTFLRSGATEVMDA 299

DB 254 LMGLPYGRTHLLKASKQNELEAQAAGPASPQKTKQAGFLAYFEICFVR- AKKRWIN 312

QY 300 SEVPYAYKANWGLVDYDKISFVKAQWLKQNNFGAMIAIDDDFTGSCDQKGFPLT 359

DB 313 QYVVPYANKKEWGVGDYDAISFGYKAFKREHFGGAMVWTLDDDFRGYFCGCTGPFPLV 372

QY 360 STLNKALGISTEGCTAPDV-----PSEPVT-----PBG 388

DB 373 HTLNLL-VNDEFSSTPSPKFWSTAVNSSRIGPEMTWTRDLTTGLGILPPG 424

RESULT 4

146470

estrogen dependent oviduct protein precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 05-Nov-1999

C:Accession: 146470

R:DeSouza, M.M.; Murray, M.K.

Endocrinology 136, 2485-2496, 1995

A:Title: An estrogen-dependent secretory protein, which shares identity with chitinases,

and embryo development.

A:Reference number: 146470; MUID:95269691; PMID:7750470

A:Accession: 146470

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-539 &lt;DES&gt;

A:Cross-references: EMBL:U16719; NID:g885600; PID:AAC48471.1; PID:g885601

Query Match 40.2%; Score 1003; DB 2; Length 539;

Best Local Similarity 50.9%; Pred. No. 5.9e-67;

Matches 192; Conservative 62; Mismatches 115; Indels 8; Gaps 4;

QY 1 YNLICVFTWQAQYRPGLSGFKPDDINPCLCTHLIYAFAGMNNETITIE-WNDVTLTKAF 59

DB 22 HKLVCTFTWAHSRPGASILPHDLDFLCTHLIFAFASNNNNQIVPKDPLDEKILYPEF 81

QY 60 NDLKRNKSLKTLIAIGGNFCTAPFTTVMVSTSONRQTFTITSVIFKFLQYFGDGLDWE 119

DB 82 NKLKERNRELKTLISIGGNFCTAPFTTVMVSTSONRQTFTITSVIFKFLQYFGDGLDWE 141

QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGGINSIQAGYEIPEL 179

DB 142 YPGLRGSPMDHRTWTFLEILLQAFKNEAQLTWPRLLLSAAVSGDPHVQKADARLL 201

QY 180 SKYLDPIHMTYDLHGSWEVGTGENSPLYKYPTETGSGNAYLNVDYVMYKKNNGAPAEKL 239

DB 202 GRLLDFISVLSDHGSWEKVTGHNPSLFLSGDPKSSA-----YAMNYWROLGVPEPEKL 256

QY 240 IVGFPEYGHFTILRNPDSNGICAPTSGDGPAGAYTRQAGFWAYEICTFLRSGATEVMDA 299

DB 257 LMGLPYGRTHLLKASKQNELEAQAAGPASPQKTKQAGFLAYFEICFVR- AKKRWIN 315

QY 300 SEVPYAYKANWGLVDYDKISFVKAQWLKQNNFGAMIAIDDDFTGSCDQKGFPLT 359

DB 316 QYVVPYANKKEWGVGDYDAISFGYKAFKREHFGGAMVWTLDDDFRGYFCGCTGPFPLA 375

QY 360 STLNKALGISTEGCTAP 376

DB 376 HTLNLL-VNDEFSSTP 391



```
RESULT 5
A49562
cartilage glycoprotein gp39 precursor - human
N:Alternate names: 39K synovial protein
C:Species: Homo sapiens (man)
C:Date: 23-Mar-1995 #sequence revision 23-Mar-1995 #text_change 22-Jun-1999
R:Accession: A49562; S10677; A33162
R:Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A:Title: Human cartilage gp-39, a major secretory product of articular chondrocytes and
A:Reference number: A49562; MUID:94064658; PMID:8245017
A:Accession: A49562
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-383 <SHA>
A:Cross-references: GB:M80927; NID:G348911; PIDN:AAAL16074.1; PID:G348912
R:Nyirkos, P.; Golde, E.E.
Biochem. J. 269, 265-268, 1990
A:Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary prote
A:Reference number: S10677; MUID:90328983; PMID:2375755
A:Accession: S10677
A:Molecule type: protein
A:Residues: 22-40, 'X', 42-45 <NY2>
C:Superfamily: Streptomyces chitinase chi40
C:Keywords: cartilage; extracellular protein; glycoprotein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-383/Product: cartilage glycoprotein gp39 #status predicted <WAT>

Query Match 39.6%; Score 987.5; DB 2; Length 383;
Best Local Similarity 50.7%; Pred. No. 5.3e-66;
Matches 186; Conservative 66; Mismatches 106; Indels 9; Gaps 5;

QY 1 YNLICYFTNWAQYRPGLSGSKFPDDINPCLCTHLIYAFAGMNNETIIEWNDVTLKAFN 60
DB 22 YKLVCYTTSWSQYREGDSCFPDADRFCLCTHIIYSFANISNDHIDTWENNDVTLGMLN 81

QY 61 DLKNRNSKLKTLAIGGWNFGTAPFTTMVSTSONRQTFTSVIKFLROYGFGDGLDWEY 120
DB 82 TLKNRNPMLKTLSSVGGWNGFSGRFSKIASNTQSRRTFKSVPPFLRTHGFGDGLDWEY 141

QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGINSIQAGYEIPELS 180
DB 142 PGR-----DKQHTTLIKEMKAEFTKEA-QPKKKQLLSAALSAGKVTIDSSYDIAKTS 195

QY 181 KYLDFHWTYDLHGSWEGYTGNSPLYKYPTETGSNAYLVNDYVMYWKNGAPAEKLI 240
DB 196 QHLDFTISIMTYDFHGAWRGTTGHSPLFRQGDASPDRESNTDYAVGYMLRLGAPASKLV 255

QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTROAGFWAYVEICTFLRSGATEWVDAS 300
DB 256 MGIPFGRSFTLAS-SETGVGAPISGPGIPGRTKEAGLAIYEICDFLR-GATVHRTLG 313

QY 301 QEVPIYAYKANEMWLDYDNISFSVKAQWLKQNNFGGAMIWALDIDFTGSCDQG-KFPILT 359
DB 314 QQVPYATKGNQWGYDDQESVSKVQYLDKROLAGAWWALDLDLDFQSGFCQDLRFPLT 373

QY 360 STLNKAL 366
DB 374 NAIKDAL 380

RESULT 6
S51327
heparin-binding glycoprotein 38K - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 19-Mar-1997 #sequence revision 19-Mar-1997 #text_change 22-Jun-1999
R:Accession: S51327
R:Shackleton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A:Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differe
A:Reference number: S51327
```

```
A:Accession: S51327
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-383 <SHA>
A:Cross-references: EMBL:Z47803; NID:G634097; PIDN:CAA87764.1; PID:G634098
C:Superfamily: Streptomyces chitinase chi40

Query Match 39.0%; Score 973.5; DB 2; Length 383;
Best Local Similarity 50.4%; Pred. No. 5.9e-65;
Matches 185; Conservative 63; Mismatches 110; Indels 9; Gaps 5;

QY 1 YNLICYFTNWAQYRPGLSGSKFPDDINPCLCTHLIYAFAGMNNETIIEWNDVTLKAFN 60
DB 22 YKLVCYTTSWSQYREGDSCFPDADRFCLCTHIIYSFANISNDHIDTWENNDVTLGMLN 81

QY 61 DLKNRNSKLKTLAIGGWNFGTAPFTTMVSTSONRQTFTSVIKFLROYGFGDGLDWEY 120
DB 82 TLKNRNPMLKTLSSVGGWNGFSGRFSKIASNTQSRRTFKSVPPFLRTHGFGDGLDWEY 141

QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGINSIQAGYEIPELS 180
DB 142 PGR-----DKRHLTTLVKEMKAEFVREALPOTE-RLLSGAVSAGKVAIDRGYDIAQIS 195

QY 181 KYLDFHWTYDLHGSWEGYTGNSPLYKYPTETGSNAYLVNDYVMYWKNGAPAEKLI 240
DB 196 QHLDFTISIMTYDFHGAWRGTTGHSPLFRQGDASPDRESNTDYAVGYMLRLGAPASKLV 255

QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTROAGFWAYVEICTFLRSGATEWVDAS 300
DB 256 MGIPFGRSFTLAS-SKTDVGAPASGPGIPGRTKEGILAIYEICDFLR-GATVRRPLG 313

QY 301 QEVPIYAYKANEMWLDYDNISFSVKAQWLKQNNFGGAMIWALDIDFTGSCDQG-KFPILT 359
DB 314 QQVPYATKGNQWGYDDQESVSKVQYLDKROLAGAWWALDLDLDFRGNFCQDLRFPLT 373

QY 360 STLNKAL 366
DB 374 SAIKDVL 380

RESULT 7
A53918
chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)
C:Species: Chelonus sp.
C:Date: 28-Jul-1995 #sequence revision 28-Jul-1995 #text_change 21-Jul-2000
C:Accession: A53918
R:Krishnan, A.; Nair, P.N.; Jones, D.
J. Biol. Chem. 269, 20971-20976, 1994
A:Title: Isolation, cloning, and characterization of new chitinase stored in active form
A:Reference number: A53918; MUID:94342256; PMID:8063715
A:Accession: A53918
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-483 <KRI>
A:Cross-references: GB:U10422; NID:G533504; PIDN:AAA61639.1; PID:G533505
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 36.1%; Score 901; DB 2; Length 483;
Best Local Similarity 38.6%; Pred. No. 2e-59;
Matches 184; Conservative 77; Mismatches 170; Indels 46; Gaps 10;

QY 3 LICYFTNWAQYRPGLSGSKFPDDINPCLCTHLIYAFAGMNNETIIE-WND----VTLYK 57
DB 24 VVCYFGASVYRQNGKFGDINGIDFTLCTHLIYFVGNGKDVKVLDPHSDLPGLNDFG 83

QY 58 AFNDLKNRNSKLKTLAIGGWNFGTAPFTTMVSTSONRQTFTSVIKFLROYGFGDGLD 117
DB 84 KFTSLRKNKPSVKIMVAVGWNAGSVPPFSQMASDQATREAFQNVVVKFLQQYQDFG 143

QY 118 WEYPSRSGSPPODKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGINSIQAGYE 177
DB 144 WEYPAQRGSGPADVKNMVKLCKALKKAFVQH-----DYILSAAVAAPETSASKSYDIA 196
```







A:Residues: 1-699 <WAT>  
A:Cross-references: GB:M57601; GB:J05599; NID:G1066341; PIDN:AAA81528.1; PID:G142688  
C:Superfamily: fibronectin type III repeat homolog  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 19.4%; Score 484; DB 2; Length 699;  
Beat Local Similarity 27.4%; Pred. No. 4.1e-28;  
Matches 130; Conservative 80; Mismatches 156; Indels 108; Gaps 18;

```
QY 1 YNLICYFTNWAQYRGLGSFKPDDINPCLCTHLIYAFAGM-----40
Db 44 YKIVGYFYSWAAYG---RNYNVADIDPKVTHINYAFADICWNGIHGNDPDSGPNPVTTWT 100

QY 41 -QNNBITTIE-----WNVDTLYKAF-----NDLKRNNSKLKTL 72
Db 101 CQNEKSQIINVENGITVLGDPWIDTG--KTFAGDTWDQPIAGNINQNLNKLKQTENLAKTI 158

QY 73 LAIGWNFGTAPFTTWSTVSQNRQFTITSVKFLROYGPDGLDLWEYFGSRG-----SP 127
Db 159 ISVGGTWSNR-FSDVAATAATREVFANSADVFLRKYNFDGVDLDWEYFVSGGLDGNKR 217

QY 128 PODKHLFTVLVKEMEEAPEQEAIESNRRLMYTAAVAGISNIQAGYEIPELSKYLDRIH 187
Db 218 PEDKQNYTLKLSKIREKLDAGAVDGKKYLLTIA--SGASATYAANTELAIAIVDWIN 275

QY 188 VMTYDLHGSGEGYTGENSEPLKYPTET-----GSNAYLNVDYVMNYWKNG 233
Db 276 IMTYDFNGAWQKISAHNAPLNVDPAASAAGVDPDANTFNVAQAQGHDA-----G 325

QY 234 APAEKLIVGFPEYHTFILRNPDSNGIGAPTSQDGDPAGAYTRQAGWAYYEI-CTFL-RS 291
Db 326 VPAAKLVGLGVFPYGRGWDGCAQAGNGQYCTCTGSSVG--TWEAGSFDYDLEANNYINKN 383

QY 292 GATEVWDASQEVPIYAYKAN--EWLGYDNIKSPSKAOWLKQNNFGGAMIWALDDFTGS 349
Db 384 GYTRVWNTAKVPYLYNASNKRKFIISYDDAESVGYKTAYIKSKGLGAMFW-----ELSG- 437

QY 350 FCDQKFLPTSTLNKALGISTEGCTAPDVPSPPVTPPGSGSGSGSGSGSGS 403
Db 438 --DRNK-----TLQNLKADLPTGGTVP-----PVDTTAPSVFGNARSTGVITANS 480
```

Search completed: June 29, 2003, 21:02:40  
Job time : 16.0414 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:55:24 ; Search time 25.4083 Seconds  
(without alignments)  
3665.470 Million cell updates/sec

Title: US-10-004-219b-9  
Perfect score: 2493  
Sequence: 1 YNLICYFTNQAQYRPLGSLF.....QQHCOAGLVFTSCNCNWP 452

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Match	DB ID	Description
1	2493	100.0	473	11 Q99PH2	Q99ph2 mus musculus
2	2488	99.8	472	11 Q9ULN1	Q9jln1 mus musculus
3	2488	99.8	473	11 Q9D803	Q9d803 mus musculus
4	2089.5	83.8	476	4 Q9BZP6	Q9bzip6 homo sapien
5	2000	80.2	365	11 Q91XA9	Q91xap6 mus musculus
6	1991.5	79.9	472	6 Q9SM17	Q9sm17 bos taurus
7	1663.5	66.7	368	4 Q9ULY4	Q9uly4 homo sapien
8	1406.5	56.4	315	4 Q9ULY3	Q9uly3 homo sapien
9	1403	56.3	262	11 Q9D7W6	Q9d7w6 mus musculus
10	1387	55.6	398	11 Q35744	O35744 mus musculus
11	1379	55.3	398	11 P70201	P70201 mus musculus
12	1376	55.2	398	11 Q91Z98	Q91z98 mus musculus
13	1373	55.1	402	11 Q8VHG1	Q8vng1 mus musculus
14	1371	55.0	488	13 Q90W34	Q90w34 bufo japoni
15	1368	54.9	402	11 Q8VH43	Q8vh43 mus musculus
16	1301.5	52.2	466	4 Q13231	Q13231 homo sapien

17	1146	46.0	387	4 Q9H3V8	Q9h3v8 homo sapien
18	1009.5	40.5	624	6 Q19118	Q19118 macaca mula
19	1008	40.4	696	5 Q9VZV2	Q9vzv2 drosophila
20	995	39.9	1013	5 Q96OM0	Q96om0 drosophila
21	985.5	39.5	383	4 Q96H17	Q96h17 homo sapien
22	976.5	39.2	457	6 Q95L83	Q95lb3 oryctolagus
23	973.5	39.0	383	6 Q29411	Q29411 sus scrofa
24	972.5	39.0	383	6 Q8SPQ0	Q8spq0 capra hitcu
25	966.5	38.8	352	11 Q9WTV1	Q9wvt1 rattus norv
26	951	38.1	260	11 Q61201	Q61201 mus musculu
27	944.5	37.9	390	4 Q96F97	Q96f97 homo sapien
28	922	37.0	396	11 Q9D7Q1	Q9d7q1 mus musculu
29	901	36.1	483	5 Q23737	Q23737 chelonus sp
30	900.5	36.1	462	5 Q8SYH0	Q8syh0 drosophila
31	892.5	35.8	381	11 Q99J84	Q99j84 mus musculu
32	879.5	35.3	460	5 Q9W2M7	Q9w2m7 drosophila
33	879	35.3	332	6 Q18949	Q18949 bos taurus
34	860.5	34.5	544	5 Q9GQC4	Q9gc4 bombyx mori
35	859.5	34.5	566	5 Q8WR52	Q8wr52 bombyx mori
36	858	34.4	553	5 P91731	P91731 hyphantria
37	857.5	34.4	543	5 Q9GV05	Q9gv05 bombyx mori
38	857.5	34.4	543	5 Q9GR93	Q9gr93 bombyx mori
39	857.5	34.4	565	5 P90710	P90710 bombyx mori
40	851	34.1	498	5 Q9W2Z3	Q9w2z3 drosophila
41	849.5	34.1	552	5 Q9GV44	Q9gv44 spodoptera
42	844.5	33.9	525	5 Q44079	Q44079 anopheles g
43	841.5	33.8	565	5 Q9GPG9	Q9gpg9 bombyx mand
44	838.5	33.6	460	5 Q8WS95	Q8ws95 glossina mo
45	826.5	33.2	520	5 Q17100	Q17100 acanthochei

## ALIGNMENTS

## RESULT 1

Q99PH2 ID: Q99PH2 PRELIMINARY; PRT; 473 AA.  
AC Q99PH2; DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Acidic mammalian chitinase (EC 3.2.1.14).  
GN CHIA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=LUNG;  
RX MEDLINE=21125893; PubMed=11085997;  
RA Boot R.G., Blommaert E.F.C., Swart E., Chauharali-Van Der Vlugt K.,  
RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;  
RT "Identification of a Novel Acidic Mammalian Chitinase Distinct from  
RT Chitotriosidase.";  
RL J. Biol. Chem. 276:6770-6778(2001).  
CC - SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
DR EMBL; AF290003; AAG60018.1; -.  
DR MGD; MGI:1932052; Chia.  
DR InterPro; IPR001579; Chitinase 18/2.  
DR InterPro; IPR002557; Chitin bind PerA.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF01607; CEM\_14; 1.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR SMART; SM00494; ChitBD2; 1.  
DR PROSITE; PS01095; CHITINASE\_18; 1.  
KW Glycosidase; Hydrolase  
SQ SEQUENCE. 473 AA; 51977 MW; 389D87557BEC0784 CRC64;  
Query Match 100.0%; Score 2493; DB 11; Length 473;  
Best Local Similarity 100.0%; Pred. No. 4.8e-171;

Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YNLICFTNWAQVRPGSGFKPDDINPCLCTHLIYAFAGMNNNEITTIENWDTLYKAFN 60  
 DB 22 YNLICFTNWAQVRPGSGFKPDDINPCLCTHLIYAFAGMNNNEITTIENWDTLYKAFN 81  
 QY 61 DLKRNRSKLTLLAIGGNWFGTAPFTTWTSTQNRQTFITSVIKFLRQYGFGLDLDEY 120  
 DB 82 DLKRNRSKLTLLAIGGNWFGTAPFTTWTSTQNRQTFITSVIKFLRQYGFGLDLDEY 141  
 QY 121 PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGINSIQAGYEIPELS 180  
 DB 142 PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGINSIQAGYEIPELS 201  
 QY 181 KYLDFHVTYDHLGSGWEGYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKLI 240  
 DB 202 KYLDFHVTYDHLGSGWEGYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKLI 261  
 QY 241 VGPEYGHFTILRNPDSNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300  
 DB 262 VGPEYGHFTILRNPDSNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 321  
 QY 301 QVVPYAYKANWLGYNISFSVKAQWLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLTS 360  
 DB 322 QVVPYAYKANWLGYNISFSVKAQWLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLTS 381  
 QY 361 TLNKGALISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGSGGFCADKADGLYPVADDR 420  
 DB 382 TLNKGALISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGSGGFCADKADGLYPVADDR 441  
 QY 421 NAFWQINGITYQOHCQAGLVFDTSCNCCNWP 452  
 DB 442 NAFWQINGITYQOHCQAGLVFDTSCNCCNWP 473

RESULT 2

Q9JLN1 PRELIMINARY; PRT; 472 AA.

ID AC Q9JLN1  
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE Putative chitinase precursor (fragment).  
 GN CHIA OR YNL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC147668; TISSUE=SKIN;  
 RA Price P.A., Harris S.C., Williamson M.K.;  
 RT "YNL, A Putative Mouse Chitinase."  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL HYDROLASES).  
 DR EMBL; AF154571; AAF31644.1; -.  
 DR MGD; MG1:1932052; Chita.  
 DR InterPro; IPR001579; Chitinase\_18/2.  
 DR InterPro; IPR002557; Chitin\_bind\_Pera.  
 DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR Pfam; PF01607; CBM 14; 1.  
 DR Pfam; PF0704; Glyco\_hydro\_18; 1.  
 DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
 DR SMART; SM00494; ChitB2; 1.  
 DR PROSITE; PS01095; CHITINASE\_18; 1.  
 KW Glycosidase; Hydrolase; Signal.  
 FT SIGNAL 1 1  
 FT CHAIN 21 472 PUTATIVE CHITINASE.  
 SQ SEQUENCE 472 AA; 51872 MW; FFF5908512C8A7F0 CRC64;

Query Match

99.8%; Score 2488; DB 11; Length 472;

Best Local Similarity 99.8%; Pred. No. 1.1e-170;  
 Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YNLICFTNWAQVRPGSGFKPDDINPCLCTHLIYAFAGMNNNEITTIENWDTLYKAFN 60  
 DB 21 YNLICFTNWAQVRPGSGFKPDDINPCLCTHLIYAFAGMNNNEITTIENWDTLYKAFN 80  
 QY 61 DLKRNRSKLTLLAIGGNWFGTAPFTTWTSTQNRQTFITSVIKFLRQYGFGLDLDEY 120  
 DB 81 DLKRNRSKLTLLAIGGNWFGTAPFTTWTSTQNRQTFITSVIKFLRQYGFGLDLDEY 140  
 QY 121 PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGINSIQAGYEIPELS 180  
 DB 141 PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGINSIQAGYEIPELS 200  
 QY 181 KYLDFHVTYDHLGSGWEGYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKLI 240  
 DB 201 KYLDFHVTYDHLGSGWEGYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKLI 260  
 QY 241 VGPEYGHFTILRNPDSNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300  
 DB 261 VGPEYGHFTILRNPDSNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 320  
 QY 301 QVVPYAYKANWLGYNISFSVKAQWLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLTS 360  
 DB 321 QVVPYAYKANWLGYNISFSVKAQWLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLTS 380  
 QY 361 TLNKGALISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGGFCADKADGLYPVADDR 420  
 DB 381 TLNKGALISTEGCTAPDVPSEPTTTPGSGSGSGSGSGSGGFCADKADGLYPVADDR 440  
 QY 421 NAFWQINGITYQOHCQAGLVFDTSCNCCNWP 452  
 DB 441 NAFWQINGITYQOHCQAGLVFDTSCNCCNWP 472

RESULT 3

Q9D803 PRELIMINARY; PRT; 473 AA.

ID AC Q9D803  
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE 220003E03Rik protein.  
 GN CHIA OR 220003E03RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itch M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Hayashizaki Y.,  
 RA Hayashaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).





```
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN 1.
SQ SEQUENCE 365 AA; 39968 MW; C900BE0C4416F1DD CRC64;

Query Match
Best Local Similarity 80.2%; Score 2000; DB 11; Length 365;
Matches 364; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 88 MVSTSNQRTFTSVIKFLROYGDFGLDLDWEYVPGSRGSPPODKHLFTVLVKEMREAFQ 147
Db 1 MVSTSNQRTFTSVIKFLROYGDFGLDLDWEYVPGSRGSPPODKHLFTVLVKEMREAFQ 60

QY 148 EATESNRPLMTAAVAGGINSIQAGVEIPELSKYLDFTIHVMYDILHGSWEGYTGNSPL 207
Db 61 EATESNRPLMTAAVAGGINSIQAGVEIPELSKYLDFTIHVMYDILHGSWEGYTGNSPL 120

QY 208 YKPTTGTGSNAYLNVDYVMNYKNGNGAPAEKLIIVGPEYGHFTILRNPSDNGIGAPTSGD 267
Db 121 YKPTTGTGSNAYLNVDYVMNYKNGNGAPAEKLIIVGPEYGHFTILRNPSDNGIGAPTSGD 180

QY 268 GPAGAYTROAGFWAYYEICTFLRSQATEVWDASQEVYAYKANENWILGYDNIKSFVKAQW 327
Db 181 GPAGAYTROAGFWAYYEICTFLRSQATEVWDASQEVYAYKANENWILGYDNIKSFVKAQW 240

QY 328 LKONNFGGAMIWALDLDFTGSCDOGKPLTSTLNKALGISTEGCTAPDVPSEPTTTP 387
Db 241 LKONNFGGAMIWALDLDFTGSCDOGKPLTSTLNKALGISTEGCTAPDVPSEPTTTP 300

QY 388 GSGSGSGSGSGSGSGSGGFCADKADGLYPVADDRNAPFQINGITYQHQHQAAGLVFTSCN 447
Db 301 GSGSGSGSGSGSGSGSGGFCADKADGLYPVADDRNAPFQINGITYQHQHQAAGLVFTSCN 360

QY 448 CCNWP 452
Db 361 CCNWP 365

RESULT 6
Q95M17 PRELIMINARY; PRT; 472 AA.
AC Q95M17;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chitin binding protein b04.
GN CBP B04.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER.
RX MEDLINE=21475601; PubMed=11591385;
RA Suzuki M., Morimatsu M., Yanashita T., Iwanaga T., Syuto B.;
RT "A novel serum chitinase that is expressed in bovine liver.";
RL FEBS Lett. 506:127-130(2001).
DR EMBL; AB051629; BAB71805.1; -.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN 1.
SQ SEQUENCE 472 AA; 52129 MW; 7A4A600E8DA04B1E CRC64;

Query Match
Best Local Similarity 79.9%; Score 1991.5; DB 6; Length 472;
Matches 348; Conservative 45; Mismatches 57; Indels 1; Gaps 1;

QY 88 MVSTSNQRTFTSVIKFLROYGDFGLDLDWEYVPGSRGSPPODKHLFTVLVKEMREAFQ 147
Db 1 MVSTSNQRTFTSVIKFLROYGDFGLDLDWEYVPGSRGSPPODKHLFTVLVKEMREAFQ 60
```

```
QY 1 YNLICVFTNWAQYRPLGSGFKFDDINPCLCTHLIYAFAGMONEITTTIENWVTLTKAFN 60
Db 22 YOLVCYFNSWAQYRPLGSGFKFDDINPCLCTHLIYAFAGMONEITTTIENWVTLTKAFN 81

QY 61 DLKRNRSKLTLLAIGGNWFGTAPFTTVMVSTSNQRTFTSVIKFLROYGDFGLDLDWEY 120
Db 82 DLKRNRSKLTLLAIGGNWFGTAPFTTVMVSTSNQRTFTSVIKFLROYGDFGLDLDWEY 141

QY 121 PGRSGSPPODKHLFTVLVKEMREAFQEAISNRRPLMTAAVAGGINSIQAGYIPELS 180
Db 142 PGRSGSPPODKHLFTVLVKEMREAFQEAISNRRPLMTAAVAGGINSIQAGYIPELS 201

QY 181 KYLDFTIHVMYDILHGSWEGYTGNSPLKYPTTGTGSNAYLNVDYVMNYKNGNGAPAEKLI 240
Db 202 KYLDFTIHVMYDILHGSWEGYTGNSPLKYPTTGTGSNAYLNVDYVMNYKNGNGAPAEKLI 261

QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTROAGFWAYYEICTFLRSQATEVWDAS 300
Db 262 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTROAGFWAYYEICTFLRSQATEVWDAS 321

QY 301 QEVYAYKANENWILGYDNIKSFVKAQWILKONNFGGAMIWALDLDFTGSCDOGKPLT 360
Db 322 QEVYAYKANENWILGYDNIKSFVKAQWILKONNFGGAMIWALDLDFTGSCDOGKPLT 381

QY 361 TLNKGALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGGFCADKADGLYPVADDR 420
Db 382 TLNKGALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGGFCADKADGLYPVADDR 440

QY 421 NAFWQINGITYQHQHQAAGLVFTSCN 451
Db 441 NAFWQINGITYQHQHQAAGLVFTSCN 471

RESULT 7
Q9ULY4 PRELIMINARY; PRT; 368 AA.
AC Q9ULY4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Novel member of chitinase family.
GN TSA1902-L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20018184; PubMed=10548734;
RA Saito A., Ozaki K., Fujiwara T., Nakamura Y., Tanigami A.;
RT "Isolation and mapping of a human lung-specific gene, TSA 1902,
RT encoding a novel chitinase family member.";
RL Gene 239:325-331(1999).
DR EMBL; AB025008; BAA86980.1; -.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN 1.
SQ SEQUENCE 368 AA; 40082 MW; 10FB970C79E19254 CRC64;

Query Match
Best Local Similarity 66.7%; Score 1663.5; DB 4; Length 368;
Matches 293; Conservative 36; Mismatches 35; Indels 3; Gaps 1;

QY 88 MVSTSNQRTFTSVIKFLROYGDFGLDLDWEYVPGSRGSPPODKHLFTVLVKEMREAFQ 147
Db 1 MVSTSNQRTFTSVIKFLROYGDFGLDLDWEYVPGSRGSPPODKHLFTVLVKEMREAFQ 60
```



```
Db 121 TAAGGAGNIOAGYBIPKSLYDLFIHVMYDYLHGSWEGYTGENSPLYKYPTETGSNAY 180
Qy 220 LNVYVYMWYKNGAPAEKLIYGVPEYHTEFLRNPDSNGIGAPTSKDGCPAGAYTRQAGF 279
Db 181 LNVYVYMWYKNGAPAEKLIYGVPEYHTEFLRNPDSNGIGAPTSKDGCPAGAYTRQAGF 240
Qy 280 WAYEYICTFLRSGATEVWDASQ 301
Db 241 WAYEYICTFLRSGATEVWDASQ 262

RESULT 10
O35744
ID AC O35744 PRELIMINARY; PRT; 398 AA.
AC O35744;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Secretory protein precursor.
GN CHI3L3 OR Ym1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264517; PubMed=11297523;
RA Chang N.C.A., Hung S.I., Hwa K.Y., Kato I., Chen J.B., Liu C.H.,
RA Chang A.C.;
RT "A macrophage protein, Ym1, transiently expressed during inflammation
RT is a novel mammalian lectin."
RL J. Biol. Chem. 276:17497-17506(2001).
DR EMBL; M94584; AAB62394.2; -
DR MGD; MGI:1330860; Chi3l3.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 398 SECRETORY PROTEIN.
FT CHAIN 22 398 SECRETORY PROTEIN.
SQ SEQUENCE 398 AA; 44458 MW; C1187661899D1D1 CRC64;

Query Match 55.6%; Score 1387; DB 11; Length 398;
Best Local Similarity 66.2%; Pred. No. 1.1e-91;
Matches 249; Conservative 55; Mismatches 72; Indels 0; Gaps 0;

Qy 1 YNLICVFTWAQYRPLGSGFKDDINPCLCTHLIYAFAGMNNETITIEWNDVTLKAFN 60
Db 22 YQLMCYTTSWAKDRPTEGSKFPGNDPCLCTHLIYAFAGMNNETITYTHEQDLRDYEALN 81

Qy 61 DLKRNKSLKTLAIGGNWFGTAPFTTWSTSONRQTFITSVIKFLROYGFGDGLDWEY 120
Db 82 GLKDNKTELKTLAIGGWKFGAPFSAMVSTFQNRQIFQSVIRFLRQYNFGLMDWQY 141

Qy 121 PGRSGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGGINSIQAGYIPELS 180
Db 142 PGRSGSPKDKHLFVLVKEMKAFEESEVEKDIPRLLITSTGAGIIDVIKSGYKIPELS 201

Qy 181 KYLDFIHVMYDYLHGSWEGYTGENSPLYKYPTETGSNAYLVNDVYMWYKNGAPAEKLI 240
Db 202 QSLDYIQVMYDYLHDPKDGTYTGENSPLYKSPYDYGKSDADLVNDSIISYWKDHGAASEKLI 261

Qy 241 VGFPEYGHTEFLRNPDSNGIGAPTSKDGCPAGAYTRQAGFWAYEICTFLRSGATEVWDAS 300
Db 262 VGFPAHGHTFLRNPDSNGIGAPTSKDGCPAGAYTRQAGFWAYEICTFLRSGATEVWDAP 321

Qy 301 QEVPPYAYKANLWGYDNISFVSKAQWLKQNNFGGAMIWAILDDPFTGSCFQCKPFLTS 360
Db 142 PGRSGSPKDKHLFVLVKEMKAFEESEVEKDIPRLLITSTGAGIIDVIKSGYKIPELS 201

Qy 181 KYLDFIHVMYDYLHGSWEGYTGENSPLYKYPTETGSNAYLVNDVYMWYKNGAPAEKLI 240
Db 202 QSLDYIQVMYDYLHDPKDGTYTGENSPLYKSPYDYGKSDADLVNDSIISYWKDHGAASEKLI 261

Qy 241 VGFPEYGHTEFLRNPDSNGIGAPTSKDGCPAGAYTRQAGFWAYEICTFLRSGATEVWDAS 300
Db 262 VGFPAHGHTFLRNPDSNGIGAPTSKDGCPAGAYTRQAGFWAYEICTFLRSGATEVWDAP 321

Qy 301 QEVPPYAYKANLWGYDNISFVSKAQWLKQNNFGGAMIWAILDDPFTGSCFQCKPFLTS 360
Db 322 QEVPPYAYQNGWYGVNVSFVKLAQWLKQNNLGGAVVWPLDMDDFSGSFCHQHFPLTS 381

Qy 361 TLKALGISTEGCTAP 376
Db 382 TLKGLDNIHSASCKGP 397
```

```
RESULT 11
P70201
ID AC P70201 PRELIMINARY; PRT; 398 AA.
AC P70201;
DT 01-FEB-1997 (TREMELrel. 02, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE ECF-L.
GN CHI3L3 OR MECF-L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BONE MARROW;
RX MEDLINE=20092902; PubMed=10625674;
RA Ohashi M., Arita H., Hayai N.;
RT "Identification of a novel eosinophil chemotactic cytokine (ECF-L) as
RT a chitinase family protein."
RL J. Biol. Chem. 275:1279-1286(2000).
DR EMBL; D87757; BAA13458.2; -
DR MGD; MGI:1330860; Chi3l3.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
SQ SEQUENCE 398 AA; 44448 MW; 9D1333FC490096793 CRC64;

Query Match 55.3%; Score 1379; DB 11; Length 398;
Best Local Similarity 66.0%; Pred. No. 4.1e-91;
Matches 248; Conservative 55; Mismatches 73; Indels 0; Gaps 0;

Qy 1 YNLICVFTWAQYRPLGSGFKDDINPCLCTHLIYAFAGMNNETITIEWNDVTLKAFN 60
Db 22 YQLMCYTTSWAKDRPTEGSKFPGNDPCLCTHLIYAFAGMNNETITYTHEQDLRDYEALN 81

Qy 61 DLKRNKSLKTLAIGGNWFGTAPFTTWSTSONRQTFITSVIKFLROYGFGDGLDWEY 120
Db 82 GLKDNKTELKTLAIGGWKFGAPFSAMVSTFQNRQIFQSVIRFLRQYNFGLMDWQY 141

Qy 121 PGRSGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGGINSIQAGYIPELS 180
Db 142 PGRSGSPKDKHLFVLVKEMKAFEESEVEKDIPRLLITSTGAGIIDVIKSGYKIPELS 201

Qy 181 KYLDFIHVMYDYLHGSWEGYTGENSPLYKYPTETGSNAYLVNDVYMWYKNGAPAEKLI 240
Db 202 QSLDYIQVMYDYLHDPKDGTYTGENSPLYKSPYDYGKSDADLVNDSIISYWKDHGAASEKLI 261

Qy 241 VGFPEYGHTEFLRNPDSNGIGAPTSKDGCPAGAYTRQAGFWAYEICTFLRSGATEVWDAS 300
Db 262 VGFPAHGHTFLRNPDSNGIGAPTSKDGCPAGAYTRQAGFWAYEICTFLRSGATEVWDAP 321

Qy 301 QEVPPYAYKANLWGYDNISFVSKAQWLKQNNFGGAMIWAILDDPFTGSCFQCKPFLTS 360
Db 322 QEVPPYAYQNGWYGVNVSFVKLAQWLKQNNLGGAVVWPLDMDDFSGSFCHQHFPLTS 381

Qy 361 TLKALGISTEGCTAP 376
Db 382 TLKGLDNIHSASCKGP 397

RESULT 12
Q91298
ID AC Q91298 PRELIMINARY; PRT; 398 AA.
AC Q91298;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Secreted protein precursor Ym2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN NCBI\_TaxID=10090;  
RP SEQUENCE FROM N.A.  
RX STRAIN=BALB/C;  
RY PubMed=11533626;  
RA Webb D.C., McKenzie A.N.J., Foster P.S.;  
RT "Expression of the Ym2 Lectin-binding Protein Is Dependent on  
RT Interleukin (IL)-4 and IL-13 Signal Transduction. IDENTIFICATION OF A  
RT NOVEL ALLERGY-ASSOCIATED PROTEIN.";  
RL J. Biol. Chem. 276:41969-41976(2001).  
DR EMBL: AV049765; AAL03953.1; -;  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
KW Signal.  
FT SIGNAL.  
SQ SEQUENCE 398 AA; 44515 MW; 79A74C45CDA03850 CRC64;  
  
Query Match 55.2%; Score 1376; DB 11; Length 398;  
Best Local Similarity 65.2%; Pred. No. 6.7e-91;  
Matches 245; Conservative 61; Mismatches 70; Indels 0; Gaps 0;  
  
QY 1 YNLICFTNWAQYRPGLSFKPDDINPCLCTHLIYAFAGMKNNEITTIWINDVTLKAFN 60  
DB 1 YNLICFTNWAQYRPGLSFKPDDINPCLCTHLIYAFAGMKNNEITTIWINDVTLKAFN 60  
QY 22 YQLMCYITSWAKDRPTEGSGFKGNIDPCLCTHLIYAFAGMKNNEITYLSEQDLRDYEALN 81  
DB 22 YQLMCYITSWAKDRPTEGSGFKGNIDPCLCTHLIYAFAGMKNNEITYLSEQDLRDYEALN 81  
QY 61 DLKRNRSKLKTLAIGGWNFGTPTTVMVSTSONRQTFTSVIKFLRQYGFGLDLDMWEY 120  
DB 61 DLKRNRSKLKTLAIGGWNFGTPTTVMVSTSONRQTFTSVIKFLRQYGFGLDLDMWEY 120  
QY 82 GLKDRNTELTLLAIGGWNFGTPTTVMVSTSONRQTFTSVIKFLRQYGFGLDLDMWEY 141  
DB 82 GLKDRNTELTLLAIGGWNFGTPTTVMVSTSONRQTFTSVIKFLRQYGFGLDLDMWEY 141  
QY 121 PGSRSGPPQDKHLFTVLKEMREAEQEAIESNRPRLMVTAAGGIGNIAGYIPELS 180  
DB 121 PGSRSGPPQDKHLFTVLKEMREAEQEAIESNRPRLMVTAAGGIGNIAGYIPELS 180  
QY 142 PGSRSGPPQDKHLFTVLKEMREAEQEAIESNRPRLMVTAAGGIGNIAGYIPELS 201  
DB 142 PGSRSGPPQDKHLFTVLKEMREAEQEAIESNRPRLMVTAAGGIGNIAGYIPELS 201  
QY 181 KYLDFIHWMTYDLHSGWEGYTGNSPLYKYPTETGSNAYLVNVDYVMYWKNGAPAEKLI 240  
DB 181 KYLDFIHWMTYDLHSGWEGYTGNSPLYKYPTETGSNAYLVNVDYVMYWKNGAPAEKLI 240  
QY 202 QSLDYIQVMTYDLHDPKNGYTGNSPLYKSPYDICKSADLNVDISIYWKDHGAASEKLI 261  
DB 202 QSLDYIQVMTYDLHDPKNGYTGNSPLYKSPYDICKSADLNVDISIYWKDHGAASEKLI 261  
QY 241 VGFPEYGHFTILRNPDSNGIGAPTSGDPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300  
DB 241 VGFPEYGHFTILRNPDSNGIGAPTSGDPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300  
QY 262 VGFPAVGHFTILSDPSKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321  
DB 262 VGFPAVGHFTILSDPSKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321  
QY 301 QBPVPYAYKANWLGVDNKSFSVKAQWLKQNNFGGAMIWAILDDPFTSGFCQGFPLTS 360  
DB 301 QBPVPYAYKANWLGVDNKSFSVKAQWLKQNNFGGAMIWAILDDPFTSGFCQGFPLTS 360  
QY 322 QBPVPYAYLGNEWGVYDNRVSPFKAKQWLKQNNLGAVVWPLDMDDESGSFCHQGRPLTT 381  
DB 322 QBPVPYAYLGNEWGVYDNRVSPFKAKQWLKQNNLGAVVWPLDMDDESGSFCHQGRPLTT 381  
QY 361 TLNKAIGISTEGCTAP 376  
DB 361 TLNKAIGISTEGCTAP 376  
QY 382 TLKRDNLNHSASCKGP 397  
DB 382 TLKRDNLNHSASCKGP 397  
  
RESULT 13  
QSVHG1 PRELIMINARY; PRT; 402 AA.  
AC QSVHG1;  
DT 01-WAR-2002 (TrEMBLrel. 20, Created)  
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE YM2.  
GN YM2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN STRAIN=129/OLA;  
RC STRAIN=129/OLA;  
RA Su W.B., Chang N.-C.A.;  
RT "Genomic organization of the ym2 gene.";  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF461142; AAL66748.1; -;  
DR InterPro; IPR001223; Glyco\_hydro\_18.

DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
SQ SEQUENCE 402 AA; 44975 MW; D935389319A59905 CRC64;  
  
Query Match 55.1%; Score 1373; DB 11; Length 402;  
Best Local Similarity 65.3%; Pred. No. 1.1e-90;  
Matches 245; Conservative 61; Mismatches 69; Indels 0; Gaps 0;  
  
QY 1 YNLICFTNWAQYRPGLSFKPDDINPCLCTHLIYAFAGMKNNEITTIWINDVTLKAFN 60  
DB 1 YNLICFTNWAQYRPGLSFKPDDINPCLCTHLIYAFAGMKNNEITTIWINDVTLKAFN 60  
QY 22 YQLMCYITSWAKDRPTEGSGFKGNIDPCLCTHLIYAFAGMKNNEITYLSEQDLRDYEALN 81  
DB 22 YQLMCYITSWAKDRPTEGSGFKGNIDPCLCTHLIYAFAGMKNNEITYLSEQDLRDYEALN 81  
QY 61 DLKRNRSKLKTLAIGGWNFGTPTTVMVSTSONRQTFTSVIKFLRQYGFGLDLDMWEY 120  
DB 61 DLKRNRSKLKTLAIGGWNFGTPTTVMVSTSONRQTFTSVIKFLRQYGFGLDLDMWEY 120  
QY 82 GLKDRNTELTLLAIGGWNFGTPTTVMVSTSONRQTFTSVIKFLRQYGFGLDLDMWEY 141  
DB 82 GLKDRNTELTLLAIGGWNFGTPTTVMVSTSONRQTFTSVIKFLRQYGFGLDLDMWEY 141  
QY 121 PGSRSGPPQDKHLFTVLKEMREAEQEAIESNRPRLMVTAAGGIGNIAGYIPELS 180  
DB 121 PGSRSGPPQDKHLFTVLKEMREAEQEAIESNRPRLMVTAAGGIGNIAGYIPELS 180  
QY 142 PGSRSGPPQDKHLFTVLKEMREAEQEAIESNRPRLMVTAAGGIGNIAGYIPELS 201  
DB 142 PGSRSGPPQDKHLFTVLKEMREAEQEAIESNRPRLMVTAAGGIGNIAGYIPELS 201  
QY 181 KYLDFIHWMTYDLHSGWEGYTGNSPLYKYPTETGSNAYLVNVDYVMYWKNGAPAEKLI 240  
DB 181 KYLDFIHWMTYDLHSGWEGYTGNSPLYKYPTETGSNAYLVNVDYVMYWKNGAPAEKLI 240  
QY 202 QSLDYIQVMTYDLHDPKNGYTGNSPLYKSPYDICKSADLNVDISIYWKDHGAASEKLI 261  
DB 202 QSLDYIQVMTYDLHDPKNGYTGNSPLYKSPYDICKSADLNVDISIYWKDHGAASEKLI 261  
QY 241 VGFPEYGHFTILRNPDSNGIGAPTSGDPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300  
DB 241 VGFPEYGHFTILRNPDSNGIGAPTSGDPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300  
QY 262 VGFPAVGHFTILSDPSKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321  
DB 262 VGFPAVGHFTILSDPSKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321  
QY 301 QBPVPYAYKANWLGVDNKSFSVKAQWLKQNNFGGAMIWAILDDPFTSGFCQGFPLTS 360  
DB 301 QBPVPYAYKANWLGVDNKSFSVKAQWLKQNNFGGAMIWAILDDPFTSGFCQGFPLTS 360  
QY 322 QBPVPYAYLGNEWGVYDNRVSPFKAKQWLKQNNLGAVVWPLDMDDESGSFCHQGRPLTT 381  
DB 322 QBPVPYAYLGNEWGVYDNRVSPFKAKQWLKQNNLGAVVWPLDMDDESGSFCHQGRPLTT 381  
QY 361 TLNKAIGISTEGCTA 375  
DB 361 TLNKAIGISTEGCTA 375  
QY 382 TLKRDNLNHSASCKA 396  
DB 382 TLKRDNLNHSASCKA 396  
  
RESULT 14  
QSVHG1 PRELIMINARY; PRT; 488 AA.  
ID QSVHG1;  
AC QSVHG1;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Toad pancreatic chitinase (EC 3.2.1.14).  
GN TPCase.  
OS Bufo japonicus (Japanese toad).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufonidae; Bufo.  
OX NCBI\_TaxID=8387;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PANCREAS;  
RA Oshima H.;  
RL Thesis (2001), Department of Department of Biology, Waseda University,  
RL Tokyo, Japan.  
DR EMBL: AJ345054; CAC87888.1; -;  
DR InterPro; IPR002086; Aldehyde\_dehydr.  
DR InterPro; IPR001579; Chitinase\_18/2.  
DR InterPro; IPR002557; Chitin bind\_PerA.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF01607; CBM 14; 1.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR PROSITE; PS00070; ALDEHYDE\_DEHYDR\_CYS; UNKNOWN\_1.  
DR PROSITE; PS01095; CHITINASE\_18; UNKNOWN\_1.  
KW Glycosidase; Hydrolase; Signal.  
FT SIGNAL.  
FT CHAIN 1 18  
FT CHAIN 19 488 TOAD PANCREATIC CHITINASE.  
SQ SEQUENCE 488 AA; 54319 MW; E75E1AD3CB4919 CRC64;  
  
Query Match 55.0%; Score 1371; DB 13; Length 488;  
Best Local Similarity 54.0%; Pred. No. 2e-90;

Matches 259; Conservative 55; Mismatches 128; Indels 38; Gaps 6;	
QY	1 YNLICVFTWAQVRPGLSFKPDINPCLCTHLIYAFAGMONEITTIEMNDVTLKAFN 60
Db	19 YKLVCIYFTWNSQTRPDQKTVFGNIDPQCTHLVAFATWNEHKIAPYEWNDVLYKQFN 78
QY	61 DLKRNRSKLTLLAIGGMNFGTAPFTTSTQNRQTFITSVIKFLRQYGFGLDLDWEY 120
Db	79 DLKQKNLVTLLAIGGMNFGTKQFTDMVASSGNRSIFIKSVIAYLRQNNFDGIDLDPEY 138
QY	121 PGRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYEIPELS 180
Db	139 PGRSGPPDKORFTYLIQEMLDAFNEEARSSGLPRLLITAAVSAGKGTIDAGYEIAKIG 198
QY	181 KYLDFTHVMTYDLHSGWEGYTGNSPLYKYPTETGSNAYLNVDYVMNYWKNNGAPAEKLI 240
Db	199 QLLDFISVMTYDFHGGWDTQSGHNSPLCKGSTDYGLQYFNHFAMNYWKNNGAPAEKLI 258
QY	241 VGFPYEGHTFILRNPSDN--GIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVWD 298
Db	259 LGFPTYGRTF--RNPNMCDVGIPVSGAGSAGPYTREAGFWAYYEICTWL--SGSTVKWI 315
QY	299 ASQEVPIYAKNEWLGYDNIKSFVSKAQLKONNFGGAMIWALDLDFTGSCDQCKFPL 358
Db	316 PDQRFYACKSNEWVGFNQESYECKVRFELKESGFGGAMVWALDLDFTGSCDQCKFPL 375
QY	359 TSTLKAIGISTEGC-----TAPD--VPSEPVTTTPPGSGSG 392
Db	376 INHLKSLLEGSTVNCPEICGGISIFTPSNTTTTTAKPDCTTPEPPTPPP----- 430
QY	393 GSGSSGSGSGFCADKADGLYEPVADNAPFQINGITYQOHCQAGLVFDTSNCNCNP 452
Db	431 --VPPVIDVDPNPFCKEKTGDLHVNPLNTNKFYICANGRTYSMKCADGLVFOASNCNCNP 488

RESULT 15

Q8VH43

PRELIMINARY; PRT; 402 AA.

AC Q8VH43

DT 01-MAR-2002 (T-EMBLrel. 20, Created)

DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)

DE Putative secretory protein.

GN YM2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=STOMACH;

RA Yang S.C., Su W.B.;

RT "Tissue-specific Expression of Ym2 Protein.";

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY065557; AAL57751.1;

DR InterPro; IPR001223; Glyco hydro 18.

DR Pfam; PF00704; Glyco\_hydro\_18; 1

DR ProDom; PD000471; Glyco\_hydro\_18; 1.

SQ SEQUENCE 402 AA; 44985 MW; D935382E89A5981A CRC64;

Query Match 54.9%; Score 1368; DB 11; Length 402;

Best Local Similarity 65.1%; Pred. No. 2.5e-90;

Matches 244; Conservative 61; Mismatches 70; Indels 0; Gaps 0;

QY	1 YNLICVFTWAQVRPGLSFKPDINPCLCTHLIYAFAGMONEITTIEMNDVTLKAFN 60
Db	22 YQLMCTYTSWAKDRPTEGSPKGNIDPCLCTHLIYAFAGMONEITLSEQDLRDYEALN 81
QY	61 DLKRNRSKLTLLAIGGMNFGTAPFTTSTQNRQTFITSVIKFLRQYGFGLDLDWEY 120
Db	82 GLKDRNTELTLLAIGKWFPGAPFFSWSTPQNRQTFIKSVIRFLRQYNFQGLNLDWQY 141
QY	121 PGRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYEIPELS 180

Db	142 PCPRGSPPKDHLFSLVQEMRKAFEEESTLNHPIRLLLTSTGAGFIDVIKSGYKIPELS 201
QY	181 KYLDFTHVMTYDLHSGWEGYTGNSPLYKYPTETGSNAYLNVDYVMNYWKNNGAPAEKLI 240
Db	202 QSLDIYQWNTYDLHDPKNGYTGNSPLYKSPYDYGKSADLNVDISIITWKDHGAASEKLI 261
QY	241 VGFPYEGHTFILRNPSDNIGIAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300
Db	262 VGFPAYGHTFILSDPSKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDPT 321
QY	301 QEVPIYAKNEWLGYDNIKSFVSKAQLKONNFGGAMIWALDLDFTGSCDQCKFPLTS 360
Db	322 QEVPIYALGNEWVGVDNVRSPKLAQLKONNFGGAVVWPLDMDDFSGSFCHQGRFPLTT 381
QY	361 TLNKAIGISTEGCTA 375
Db	382 TLKRDNLNVHSASCKA 396

Search completed: June 29, 2003, 21:01:46  
Job time : 28.4083 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:51:08 ; Search time 8.02367 seconds  
(without alignments)  
2336.500 Million cell updates/sec

Title: US-10-004-219b-9

Perfect score: 2493

Sequence: 1 YNLICYFTNQAQYRPLGSGF.....QQHCQAGLVFDTSCNCCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1051	42.2	721	1 OGP MOUSE	Q62010 mus musculus
2	1042.5	41.8	671	1 OGP MESAU	Q60557 mesocricetu
3	1013	40.6	678	1 OGP HUMAN	Q12889 homo sapien
4	1010	40.5	537	1 OGP BOVIN	Q28042 bos taurus
5	1009.5	40.5	623	1 OGP PAPAN	P36718 papio anubi
6	1003	40.2	539	1 OGP SHEEP	Q28542 ovis aries
7	990.5	39.7	527	1 OGP PIG	Q28990 sus scrofa
8	987.5	39.6	383	1 C3L1 HUMAN	P36222 homo sapien
9	948.5	38.0	390	1 C3L2 HUMAN	Q15782 homo sapien
10	892.5	35.8	381	1 C3L1 MOUSE	Q61362 mus musculu
11	889.5	35.7	617	1 CHIT_CABEL	Q11174 caenorhabdi
12	863.5	34.6	554	1 CHIT_MANSE	P36362 manduca sex
13	810	32.5	504	1 CHIT BRUMA	P29030 brugia mala
14	484	19.4	699	1 CH11 BACCI	P20533 bacillus ci
15	466.5	18.7	423	1 CH11 APHAL	P32470 aphanocla
16	442.5	17.7	423	1 CH14 TRIHA	P48827 trichoderm
17	422	17.0	427	1 CH11_COCIM	P54196 coccidioid
18	365.5	14.7	499	1 CH1B_SERMA	P11797 serratia ma
19	351.5	14.1	563	1 CH1A_SERMA	P07254 serratia ma
20	341.5	13.7	1046	1 CHIT VIBFU	P06156 vibrio furn
21	328	13.2	550	1 CHIT NPVOF	Q10363 orgyia pseu
22	327.5	13.1	619	1 CHIT_STRLI	P36909 streptomyc
23	325.5	13.1	820	1 CH1A_ALTSO	P32823 streptomys
24	315	12.6	610	1 CHIT_STRPL	P11220 streptomys
25	300	12.0	551	1 CHIT NPVAC	P14684 autographa
26	233.5	9.4	1146	1 KTXA_KIULA	P09805 kluyveromyc
27	197.5	7.9	385	1 DIAC_HUMAN	Q01450 homo sapien
28	185.5	7.4	367	1 DIAC_RAT	Q01460 rattus norv
29	169.5	6.8	427	1 YAAH_BACSU	P37531 bacillus su
30	150	6.0	524	1 CH1D_BACCI	P27050 bacillus ci
31	136	5.5	491	1 YKCS_CABEL	P41996 caenorhabdi
32	131	5.3	597	1 CH1X_STROI	Q05638 streptomyc
33	130.5	5.2	812	1 PABD_ECOLI	P06970 escherichia

## ALIGNMENTS

### RESULT 1

OGP_MOUSE	STANDARD;	PRT;	721 AA.
AC	Q62010;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)		
DE	(Oviductin) (Estrogen-dependent oviduct protein).		
GN	OVGP1 OR OGP OR CHIT5.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ICR; TISSUE=Oviduct;		
RX	MEDLINE=96115001; PubMed=7492880;		
RA	Sendai Y., Komiya H., Suzuki K., Onuma T., Kikuchi M., Hoshi H.,		
RA	Araki Y.;		
RT	"Molecular cloning and characterization of a mouse oviduct-specific		
RT	glycoprotein."		
RL	Biol. Reprod. 53:285-294(1995).		
CC	-!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE		
CC	IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.		
CC	-!- SUBCELLULAR LOCATION: Secretory granules.		
CC	-!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF THE OVIDUCT.		
CC	-!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; D32137; BAA06863.1; .		
DR	MGI; MGI:106661; Ovgnl.		
DR	InterPro; IPR001579; Chitinase 18/2.		
DR	InterPro; IPR001223; Glyco_hydro_18.		
DR	Pfam; PF00704; Glyco_hydro_18; 1.		
DR	ProDom; PD000471; Glyco_hydro_18; 1.		
DR	PROSITE; PS01095; CHITINASE 18; FALSE NEG.		
KW	Glycoprotein; Fertilization; Repeat; Signal.		
FT	SIGNAL 1 21 POTENTIAL.		
FT	CHAIN 22 721 OVIDUCT-SPECIFIC GLYCOPROTEIN. S-K-T-T-		
FT	DOMAIN 486 632 21 X 7 AA TANDEM REPEATS OF S-K-T-T-		
FT	[TAP]-G-[IV].		
FT	CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL) .		
FT	CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL) .		
FT	CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL) .		
SQ	SEQUENCE 721 AA; 78807 MW; 37246C8F01665652 CRC64;		

Query Match 42.2%; Score 1051; DB 1; Length 721;  
Best Local Similarity 52.0%; Pred. No. 8.4e-68;

```
Matches 206; Conservative 61; Mismatches 119; Indels 10; Gaps 6;
QY 1 YNLICFTNWAQYRPGSGFKPDDINPCLCTHLIYAFAGQNNNEITTIENWD-VLYKAF 59
DB 22 YKLVCFYTNWAHSRPGSPASIMPHDLDFLCTHLIFAFASNNQIVAKNLDQENLVYPEF 81
QY 60 NDLKRNKSLKTLALGCGNFGTAPFTTWSTSONQRTFTITSVKEFLRQVGFGLDLWE 119
DB 82 NKLKERNRELKTLISLGGNFGTSRTMTSLANREKIDSVISFLRHGDFGLDLFL 141
QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRPRLMVTAAVAGGINSIQAGYEIPEL 179
DB 142 YPGLRGSPHPRNFWFLIEELQFAFERALLTQHPRLLLSAAVSGIPSIHTSYDALL 201
QY 180 SKYLDHIVMTYDLHGSWEYTGENSEPLKYPTETGSNAYLVNDVYNNYKNGAPAEKL 239
DB 202 GRRLDPINVLSDYHGSWEKFTGHNSPLSLPDSKSSA-----YANNWRKLGTPADKL 256
QY 240 IVGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYEICTLRSGATEVWDA 299
DB 257 IMGFTYGRNFYLLKESKGLQTASGPGSPGKYTKQAGFLAYEYVCSFVQR-AKGHWD 315
QY 300 SQBPVYAYKANEWLGYNDYNTKFSVKAQMLKONNFGGAMIWDLDFTSFCQDGKPLT 359
DB 316 YQVYVAFKGEWLGYDDTISFSYKAMYKREHFGGAMVWTLDMDDVRGTFCGNGPPLV 375
QY 360 STINKALGISTEGCTAPDPSEPVTPPGSGSGGS 395
DB 376 HILNELL-VOTESNSTP-LPQFWFTSSV-NASPGS 408

RESULT 2
ID_OGP_MESAU STANDARD; PRT; 671 AA.
AC Q60527; Q60526;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (2P-0).
GN OVGPI OR OGP.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Oviduct;
RX MEDLINE=96115007; PubMed=7492686;
RA Suzuki K., Sendai Y., Osuma T., Hoshi H., Hiroi M., Araki Y.;
RT "Molecular characterization of a hamster oviduct-specific
RT glycoprotein.";
RL Biol. Reprod. 53:345-354 (1995).
RN [2]
RP SEQUENCE OF 22-671 FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=96192955; PubMed=8607967;
RA Paquette Y., Merlen Y., Malette B., Bleau G.;
RT "Allelic polymorphism in the hamster oviductin gene is due to a
RT variable number of mucin-like tandem repeats.";
RL Mol. Reprod. Dev. 42:388-396 (1995).
RN [3]
RP SEQUENCE OF 14-671 FROM N.A., AND REVISIONS.
RC TISSUE=Oviduct;
RA Paquette Y.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 22-39.
RX MEDLINE=94058981; PubMed=8240241;
RA Malette B., Bleau G.;
RT "Biochemical characterization of hamster oviductin as a sulphated
RT zona pellucida-binding glycoprotein.";
```

```
RL Biochem. J. 295:437-445 (1993).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC MIGHT ACT AS A PROTECTIVE SECRETION INFLUENCING THE FIRST STEPS OF
CC THE REPRODUCTIVE PROCESS NECESSARY FOR THE NORMAL TRIGGERING OF
CC FERTILIZATION AND EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- PTM: HIGHLY O-GLYCOSYLATED AND ALSO N-GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; D32218; BAA06977.1; -.
CC EMBL; U15048; AAC53584.1; -.
CC InterPro; IPR001579; Chitinase_18/2.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE NEG.
CC Glycoprotein; Fertilization; Repeat; Signal.
CC SIGNAL 1 21
CC FT CHAIN 22 671 OVIDUCT-SPECIFIC GLYCOPROTEIN.
CC FT DOMAIN 490 609 8 X 15 AA TANDDEM REPEATS.
CC FT REPEAT 490 504 1.
CC FT REPEAT 505 519 2.
CC FT REPEAT 520 534 3.
CC FT REPEAT 535 549 4.
CC FT REPEAT 550 564 5.
CC FT REPEAT 565 579 6.
CC FT REPEAT 580 594 7.
CC FT REPEAT 595 609 8.
CC FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 511 511 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 526 526 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 541 541 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CONFLICT 26 26 C -> A (IN REF. 4).
CC FT CONFLICT 33 33 H -> I (IN REF. 4).
CC FT CONFLICT 137 137 D -> G (IN REF. 2 AND 3).
CC FT CONFLICT 153 153 R -> Q (IN REF. 2 AND 3).
CC FT CONFLICT 165 165 F -> Y (IN REF. 2 AND 3).
CC FT CONFLICT 193 193 Q -> L (IN REF. 2 AND 3).
CC FT CONFLICT 531 545 MISSING (IN REF. 2 AND 3).
CC FT CONFLICT 595 595 T -> I (IN REF. 2 AND 3).
CC SQ SEQUENCE 671 AA; 73250 MW; BB57E0E514EC1972 CRC64;

Query Match 41.8%; Score 1042.5; DB 1; Length 671;
Best Local Similarity 51.4%; Pred No. 3 1e-67;
Matches 204; Conservative 62; Mismatches 122; Indels 9; Gaps 5;
QY 1 YNLICFTNWAQYRPGSGFKPDDINPCLCTHLIYAFAGQNNNEITTIENWD-VLYKAF 59
DB 22 YKLVCFYTNWAHSRPGSPASIMPHDLDFLCTHLIFAFASNNQIVAKNLDQENLVYPEF 81
QY 60 NDLKRNKSLKTLALGCGNFGTAPFTTWSTSONQRTFTITSVKEFLRQVGFGLDLWE 119
DB 82 NKLKERNRELKTLISLGGNFGTSRTMTSLANREKIDSVISFLRHGDFGLDLFL 141
QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRPRLMVTAAVAGGINSIQAGYEIPEL 179
DB 142 YPGLRGSPHPRNFWFLIEELQFAFERALLTQHPRLLLSAAVSGIPSIHTSYDALL 201
QY 180 SKYLDHIVMTYDLHGSWEYTGENSEPLKYPTETGSNAYLVNDVYNNYKNGAPAEKL 239
```



Db 202 GRRLDIFINVSLDHLGWSKSTCHNSPLSPEDPKSSA-----FAMNWRNLGAPADKL 256  
 QY 240 IVGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEWDA 299  
 Db 257 LMGFPAYGRTFHLRLSKNGLSQASNGPASPCKYTKQAGFLAYEVCSPQR-AEKHWID 315  
 QY 300 SEQVPYAYKANLWGLYDNTKTSVKAQWLKQNNFGGAMTWAIDLDFTGSPCDQGGKFPJT 359  
 Db 316 HQVVPYAYKGEWGVYDDAVSYKAMFVKKEHFGGAMVWTLMDMDVTRGTCGNGPFPPLV 375  
 QY 360 STLNKALGISTEGCTAPDVPSEPVTPPGSGGGSS 396  
 Db 376 HILNELL-VRAEFNSTP-LPQFWFTLPVNSGPGSES 410

## RESULT 3

OGP\_HUMAN STANDARD; PRT; 678 AA.  
 AC Q12889; Q15841;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
 DE (Oviductin) (estrogen-dependent oviduct protein).  
 GN OVGPI OR OGP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oviduct;  
 RA MEDLINE=95119256; PubMed=7819450;  
 RA Arias E.B., Verhage H.G., Jaffe R.C.;  
 RT "Complementary deoxyribonucleic acid cloning and molecular  
 RT characterization of an estrogen-dependent human oviductal  
 RT glycoprotein";  
 RL Biol. Reprod. 51:685-694 (1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Jaffe R.C.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
 CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.  
 CC -!- SUBCELLULAR LOCATION: Secretory granules.  
 CC -!- TISSUE SPECIFICITY: OVIDUCT.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U09550; AAB46946.1;  
 EMBL; U58010; AAB04126.1;  
 EMBL; U58001; AAB04126.1; JOINED.  
 EMBL; U58002; AAB04126.1; JOINED.  
 EMBL; U58003; AAB04126.1; JOINED.  
 EMBL; U58004; AAB04126.1; JOINED.  
 EMBL; U58005; AAB04126.1; JOINED.  
 EMBL; U58006; AAB04126.1; JOINED.  
 EMBL; U58007; AAB04126.1; JOINED.  
 EMBL; U58008; AAB04126.1; JOINED.  
 EMBL; U58009; AAB04126.1; JOINED.  
 Genew; HGNC:8524; OVGPI.  
 MIM; 603578;  
 InterPro; IPR001579; Chitinase\_18/2.  
 DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
 DR ProDom; PD000471; Glyco\_hydro\_18; 1.

DR PROSITE; PS01095; CHITINASE 18; FALSE\_NEG.  
 KW Glycoprotein; Fertilization; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 678  
 FT CARBOHYD 402 402  
 FT CARBOHYD 441 441  
 FT CARBOHYD 580 580  
 FT CARBOHYD 596 596  
 FT CARBOHYD 648 648  
 FT CONFLICT 477 477  
 FT CONFLICT 511 511  
 FT CONFLICT 514 514  
 FT CONFLICT 676 676  
 SQ SEQUENCE 678 AA; 75421 MW; 245F2CEDCE92768B CRC64;  
 Query Match 40.6%; Score 1013; DB 1; Length 678;  
 Best Local Similarity 48.8%; Pred. No. 4,1e-65;  
 Matches 203; Conservative 58; Mismatches 125; Indels 30; Gaps 5;  
 QY 1 YNLICYFTNWAQYRGLGSKFKPDDINPCLCTHLIYAFAGMQNNEITTIEMND-VTLTKAF 59  
 Db 22 HKLVCFYFTNWAHSRPGASILLPHDLDPFLCTHLIFAFASMNQIVAKDLQDEKILYPEF 81  
 QY 60 NDLKVRNSKLTLLAIGWNFGTAPFTTMVSTSQNRQFTITSVIKFLQYGGDGLDWE 119  
 Db 82 NKLKERNRELKTLISIGGNFCTSFRTTLMSTFANREKFIASVISLLRTHDFDGLDLPFL 141  
 QY 120 YGSGSGSPQDKHLETVLVKEMREAFQEAIESNPRLMTVAAGVGSINIQAGYEIPEL 179  
 Db 142 YFGLRGSPMDHRTWFLFLEILLFAFRKEALLTMRPLLSSAAVSGVPHIVQTSVDVREL 201  
 QY 180 SKYLDPIHVMYDHLGWSWEGYTGNSPLYKYKPYTETGNSAYLNVDMYVKNKGAPAEKL 239  
 Db 202 GLLDFINVLSDHLGSWERFTGHSPLFSLPEPKSSA-----YAMNWRKLGAPSEKL 256  
 QY 240 IVGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEWDA 299  
 Db 257 IMGPIYGRTRFLRLKASKNGLQARAIGPASPGYTKQEGFLAYFEICSPV-WGAKKHGWD 315  
 QY 300 SEQVPYAYKANLWGLYDNTKTSVKAQWLKQNNFGGAMTWAIDLDFTGSPCDQGGKFPJT 359  
 Db 316 HQVVPYAYKGEWGVYDDAVSYKAMFVKKEHFGGAMVWTLMDMDVTRGTCGNGPFPPLV 375  
 QY 360 STLNKALG-----ISTEGCTAPDVPSEPVTP-----PPSGSGS 392  
 Db 376 YVNDILVRAEFSSTSLPQFWLSSAVNSSTDPERLAVTTATWTTDSKILPFGGEAG 431

## RESULT 4

OGP\_BOVIN STANDARD; PRT; 537 AA.  
 AC Q28042;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
 DE (Oviductin) (Estrogen-dependent oviduct protein) (Fragment).  
 GN OVGPI OR OGP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-47.  
 RC TISSUE=Oviduct;  
 RX MEDLINE=94257768; PubMed=8199272;  
 RA Sendai Y., Abe H., Kikuchi M., Satoh T., Hoshi H.;  
 RT "Purification and molecular cloning of bovine oviduct-specific  
 RT glycoprotein";  
 RL Biol. Reprod. 50:927-934 (1994).  
 CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
 CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.

```
CC -1- SUBCELLULAR LOCATION: Secretory granules.
CC -1- TISSUE SPECIFICITY: OVIDUCT.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D16639; BAA04065.1; -
CC InterPro; IPR001579; Chitinase 18/2.
CC Pfam; PF00704; Glyco_hydro_18; 1-
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT NON TER 1 1
FT CHAIN <1 18 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 399 399 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SEQUENCE 537 AA; 59617 MW; 59617 MW; CFCDEE6F0212D791 CRC64;
CC -----
Query Match 40.5%; Score 1010; DB 1; Length 537;
Best Local Similarity 49.2%; Pred. No. 4.9e-65;
Matches 203; Conservative 53; Mismatches 125; Indels 32; Gaps 6;
CC -----
QY 1 YNLICFTNWAQYRGLSGSPKDDINPCLCTHLYAFAGMNNQITTEIWNND-VTLTKAP 59
DB 19 HKLVCTFTNWAQYRGLSGSPKDDINPCLCTHLYAFAGMNNQITTEIWNND-VTLTKAP 78
QY 60 NDLKRNNSKLTKLLAIGCNWFGTAPFTTVMSTSONRQFTITSVTKELQYQFGDGLDWE 119
DB 79 NDLKRNNSKLTKLLAIGCNWFGTAPFTTVMSTSONRQFTITSVTKELQYQFGDGLDWE 138
QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRLMTVAAGVAGGINSIQAGYIPEL 179
DB 139 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRLMTVAAGVAGGINSIQAGYIPEL 198
QY 180 SKYLDFIHVMYDLHGSWEQYTGNSPLYKPYTGTGNSAYLVNVDYVMYKNGNAPAEKL 239
DB 199 GRLLDFISVLSDYHLSWEKVTGNSPLFLPDGPKSSA-----YANNYWRKLGAPSEKL 253
QY 240 IVGPEYGHFTILRNPDSNGIGAPTSGDGPAGATYQAGFWAYYEICTFLRSAGTEVWDA 299
DB 254 LMGLPTYGRTHLTKASQNELRAQAGVSPGKYTKQAGFLAYEICTFLRSAGTEVWDA 312
QY 300 SQEVPYAYKANWGLYDNIKSFVKAQWLKQNNFGAMTWAIDLDDFTGSCDQGFPLT 359
DB 313 DQVVPYAFKGEWGVDDAISFGYKAFKIKREHFGAMWTLDDDFRGYFCGTGFPFLV 372
QY 360 STLNKALGISTECTAPDV-----PSEPVT-----PPG 388
DB 373 HTLNLL-VNDEPSTSPKFWSTAVNSSRIGPEMTWTTRDITGLGILPPG 424
RESULT 5
OGP PAPAN STANDARD; PRT; 623 AA.
AC P36718;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein).
GN OVGPI OR OGP.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OC NCBI_TaxID=9555;
```

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=98244335; PubMed=9584944;
RA Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,
RA Donnelly K.M., Arias E.B., Jaffe R.C.;
RT "The baboon oviduct: characteristics of an oestradiol-dependent
RT oviduct-specific glycoprotein.";
RL Hum. Reprod. Update 3:541-552(1997).
RN (2)
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=91367180; PubMed=1716345;
RA Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,
RA Jaffe R.C.;
RT "Cloning of a recombinant complementary DNA to a baboon (Papio
RT anubis) estradiol-dependent oviduct-specific glycoprotein.";
RL Mol. Endocrinol. 5:356-364(1991).
CC -1- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Secretory granules.
CC -1- TISSUE SPECIFICITY: OVIDUCT.
CC -1- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59903; AAB39765.1; -
CC FIR; A37954; A37954.
CC InterPro; IPR001579; Chitinase 18/2.
CC Pfam; PF00704; Glyco_hydro_18; 1-
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 623 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SEQUENCE 623 AA; 69291 MW; 9E21CE481FFP1268 CRC64;
CC -----
Query Match 40.5%; Score 1009.5; DB 1; Length 623;
Best Local Similarity 52.6%; Pred. No. 6.5e-65;
Matches 193; Conservative 57; Mismatches 110; Indels 7; Gaps 3;
CC -----
QY 1 YNLICFTNWAQYRGLSGSPKDDINPCLCTHLYAFAGMNNQITTEIWNND-VTLTKAP 59
DB 22 HKLVCTFTNWAQYRGLSGSPKDDINPCLCTHLYAFAGMNNQITTEIWNND-VTLTKAP 81
QY 60 NDLKRNNSKLTKLLAIGCNWFGTAPFTTVMSTSONRQFTITSVTKELQYQFGDGLDWE 119
DB 82 NDLKRNNSKLTKLLAIGCNWFGTAPFTTVMSTSONRQFTITSVTKELQYQFGDGLDWE 141
QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRLMTVAAGVAGGINSIQAGYIPEL 179
DB 142 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRLMTVAAGVAGGINSIQAGYIPEL 201
QY 180 SKYLDFIHVMYDLHGSWEQYTGNSPLYKPYTGTGNSAYLVNVDYVMYKNGNAPAEKL 239
DB 202 GRLLDFISVLSDYHLSWEKVTGNSPLFLPDGPKSSA-----YANNYWRKLGAPSEKL 256
QY 240 IVGPEYGHFTILRNPDSNGIGAPTSGDGPAGATYQAGFWAYYEICTFLRSAGTEVWDA 299
DB 257 INGIPYGRTHLTKASQNELRAQAGVSPGKYTKQAGFLAYEICTFLRSAGTEVWDA 315
QY 300 SQEVPYAYKANWGLYDNIKSFVKAQWLKQNNFGAMTWAIDLDDFTGSCDQGFPLT 359
DB 313 DQVVPYAFKGEWGVDDAISFGYKAFKIKREHFGAMWTLDDDFRGYFCGTGFPFLV 372
```





Db 196 QHLDLISINTYDFHGAMRQTTGHSPLFRGQDASPDRESNTDYAVGVNMLRGAPASKLV 255  
QY 241 VGPBYGHFTFILRNPDSNDGIGAPTSQDGPAGAYTQAGFWAYYEICTFLRSGATEVMDAS 300  
Db 256 MGIPFGSFTLAS-SETGVGAPISGPIGPFRTKEAGTFLAYEICDFLR-CATVHRTIG 313  
QY 301 QVVPYAYKANWGLVDNKSFSVKAQWLKQNNFGGAMIWALDLDFTSGFCDOG-KPPLT 359  
Db 314 QQVPYATKGNQWGVDDQDSVSKVQYLKDRQLAGAMVWALDLDFTSGFCDOGLRFP 373  
QY 360 STLKAL 366  
Db 374 NAKDAL 380  
RESULT 9  
C3L2 HUMAN  
ID C3L2 HUMAN STANDARD; PRT; 390 AA.  
AC Q15782; Q15783; Q15749;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein 39).  
GN Chi3L2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).  
RA Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R., Mak T.W.;  
RT "Cloning of a novel lymphoid restricted human chitinase and localization to lp13.3";  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
SEQUENCE FROM N.A. (SHORT ISOFORM).  
RC TISSUE=Articular cartilage;  
RX MEDLINE=96325055; PubMed=8702629;  
RA Hu B., Trinh K., Figueira W.F., Price P.A.;  
RT "Isolation and sequence of a novel human chondrocyte protein related to mammalian members of the chitinase protein family";  
RL J. Biol. Chem. 271:19415-19420(1996).  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHOWN HERE) AND A LONG FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING  
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED BY SYNIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN, PANCREAS, AND LIVER.  
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U58515; AAB04534.1; -  
CC EMBL; U58514; AAB04533.1; -  
CC EMBL; U49835; AAC50597.1; ALT\_INIT.  
CC Genew; HGNC:1933; CHI3L2.  
CC MIM; 601526; -  
CC InterPro; IPR001579; Chitinase\_18/2.  
CC InterPro; IPR001223; Glyco\_hydro\_18.  
CC Pfam; PF00704; Glyco\_hydro\_18; 1.  
CC ProDom; PD000471; Glyco\_hydro\_18; 1.  
CC PROSITE; PS01095; CHITINASE\_18; FALSE NEG.  
KW Hydrolase; Signal; Alternative splicing.  
FT SIGNAL 1 26 POTENTIAL.  
FT CHAIN 27 390 CHITINASE 3-LIKE PROTEIN 2.  
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).

FT VARSPLIC 1 12 MGATTMDQKSLW -> SQQMLWGSDEVLVLCVCHTHIFT  
FT AIWDSRTFGSTDFSPHSI (IN LONG ISOFORM).  
SQ SEQUENCE 390 AA; 43500 MW; 97B86A2F3AA35677 CRC64;  
Query Match 38.0%; Score 948.5; DB 1; Length 390;  
Best Local Similarity 47.7%; Pred. No. 8.2e-61;  
Matches 176; Conservative 76; Mismatches 108; Indels 9; Gaps 5;  
QY 1 YNLICYFTNWAQYRGLGSKFKPDIDNPLCLTHLYAFAGMNMNEITTIEMNDVTLKAFN 60  
Db 27 YKLVCYFTNWSQDRQEPCKFTPENIDPFLCSHLIYVSFASIENNKVIKDKSEVMYQRTIN 86  
QY 61 DLKRNNSKLTLAIGGNWPGTAPTTWVSTSONRQTFTITSVIKFLROYGDFGLDMEY 120  
Db 87 SLKTKNPKLKILLSIGGYLFGSKGFHPMVDSSSTRLEFINSLILFRNHNFDGLDVSIMY 146  
QY 121 PGRSGPPQDKHLFTVLVKEMREAFQEALESNRPRLMVTAAVAGGISNIQAGYIPELS 180  
Db 147 PDQK----ENTH-FTVLIHELAEPQKDFTKSTKERLLLTAGVSAGROMINDSYQVEKLA 201  
QY 181 KYLDFIHVMTYDLHGSEWG--YTGENSPLYKYPTETGNSAVLVNDVYVMYKNGNAPAEK 238  
Db 202 KDLDFINLLSDFHGSWEKPLITGHNSPLSKGWQDRGPSSYNNVEYAVGYMTHKMPSEK 261  
QY 239 LIVGPYGHFTILRNPDSNDGIGAPTSQDGPAGAYTQAGFWAYYEICTFLRSGATEVMD 298  
Db 262 VVMGIPTVGHSTLAS-AETTVGAPASGPAAGPITESGFLAYYEICQFLK-GAKITRL 319  
QY 299 ASQEVYAYKANWGLVDNKSFSVKAQWLKQNNFGGAMIWALDLDFTSGFCDOGKPL 358  
Db 320 QQQVYPAVKGNQWGVDDVKSMTETKVQFLKNLGLGAMISIDMDFTGKSCNQGPVL 379  
QY 359 TSTLNKALG 367  
Db 380 VQAVKRSLG 388  
RESULT 10  
C3L1 MOUSE  
ID C3L1 MOUSE STANDARD; PRT; 381 AA.  
AC Q61362;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-39) (BRP39 protein).  
GN CHI3L1 OR BRP39.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Breast;  
RX MEDLINE=95060797; PubMed=7970700;  
RA Morrison B.W., Leder P.;  
RT "neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors";  
RL Oncogene 9:3417-3426(1994).  
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Extracellular (By similarity).  
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).







```
FT SIGNAL 1 22
FT CHAIN 23 504 ENDOCHITINASE.
FT DOMAIN 23 400 CATALYTIC.
FT DOMAIN 401 450 SER/THR-RICH (LINKER).
FT DOMAIN 407 448 3 X 14 AA APPROXIMATE TANDEM REPEATS.
FT ACT SITE 148 148 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 504 AA; 55971 MW; A78BE7BFBE3709B CRC64;

Query Match 32.5%; Score 810; DB 1; Length 504;
Best Local Similarity 35.2%; Pred. No. 9.5e-51;
Matches 168; Conservative 71; Mismatches 198; Indels 40; Gaps 6;

QY 5 CYPTNWAQYRGLSGFKPDINPCLCTHLYAFAGMGN-NEITTIENNDVT-----LYK 57
DB 27 CYVTNWAQYRGLSGFKPLGNIPNGLCTHLYAFAGMDELGSKPFENWDETEWSKGMYS 86

QY 58 AFNDLKNRNSKLTLAIGAGNFGTAPPTMTWTSQNRQTFTSVTKFLROYGFDGLDLD 117
DB 87 AVTKLRETNPGLKVLSSYGGYNFSGSAIPTGIAKSAQKTERFIKSAIAFLRKNFDFGLD 146

QY 118 WEYVPGSGSPPODKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYEIP 177
DB 147 WEYVPG-----VAEEH--AKLVEAMKTAFAVEEAKTSKQRLLLTAASAGKGTIDGSYVE 200

QY 178 ELSKYLDIFHVMTYDLHGSWEGYTGNSPLYKYPTETSGSNAYLVNDVVMYKNNNGAPAE 237
DB 201 SLGKNFDLLFLMSYDLHGSWEKNVDLHGLHPTKGEVSGIGIFNTFEAADYWSKGMPE 260

QY 238 KLVGPEYGHFTILRPSDNGIGATSGDPAGATROAGFWAYEICTFLRSGATFW 297
DB 261 KIIIGIPMTAQGWTLDPNPSETAIGAAASRPSSASKTNPAGGTASYWEICKYLKGGKETV 320

QY 298 DASQEVYAYKANWELGYDNISFSVKAOWLKONNFGGAMIWALDDFTGSGCDGKFP 357
DB 321 HQEGVCAYVWGDQWGYNEETIRIMKWLKEKYGGAFIWALDDFTGSGCGKGP 380

QY 358 LSTLNKALGISTEGCTAPDVPSEPTVTPPGSGSGSGSGSGSGSGGFCADKA----- 410
DB 381 LLNAISSELEGESN---PEITTEPSITETEAYETDEETSETAYOTDEETSET 437

QY 411 -----DGLYPVADRNAPWOCINGITTYQHCQAGLVPDTSNCNCN 450
DB 438 ATTYPDTEGQECRPERDGLFPHPTDCHLFIOCANNIAYVMQCPATTFNFDAIKVCD 494

RESULT 14
CH11_BACCI STANDARD; PRT; 699 AA.
AC P20533;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Chitinase A1 precursor (EC 3.2.1.14).
GN CH1A1.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1397;
RN [1]_TaxID=1397;
RP SEQUENCE FROM N.A.
RX STRAIN=WL-12; PubMed=2203782;
RX MEDLINE=90368776; PubMed=2203782;
RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;
RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed
RT its evolutionary relationship to Serratia chitinase and to the type
RT III homology units of fibronectin.";
RL J. Biol. Chem. 265:15659-15665(1990).
RN [2]
RP MUTAGENESIS.
RX STRAIN=WL-12;
RX MEDLINE=93366760; PubMed=8103047;
RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,
RA Uchida M., Tanaka H.;
RT "Identification of glutamic acid 204 and aspartic acid 200 in in
```

```
RT chitinase A1 of Bacillus circulans WL-12 as essential residues for
RT chitinase activity.";
RL J. Biol. Chem. 268:18567-18572(1993).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M57601; AAA81528.1; -.
CC PIR; A38368; A38368.
CC HSP; P02751; 1FNA.
CC InterPro; IPR003610; CBM 5.12.
CC InterPro; IPR001579; Chitinase_18/2.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR003962; FNIII_repeat.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00041; fn3; 2.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC Pfam; PF02839; CBM 5.12; 1.
CC PRINTS; PR00014; FNTYPEIII.
CC ProDom; PD000471; Glyco_hydro_18; 2.
CC SMART; SM00495; ChtBD3; 1.
CC SMART; SM00060; FN3; 2.
CC PROSITE; PS01095; CHITINASE 18; 1.
CC Hydrolase; Glycosidase; Chitin degradation; Signal; Repeat.
FT SIGNAL 1 41
FT CHAIN 42 699 CHITINASE A1.
FT DOMAIN 42 460 CATALYTIC.
FT DOMAIN 465 549 FIBRONECTIN TYPE-III (R-1).
FT DOMAIN 560 644 FIBRONECTIN TYPE-III (R-2).
FT ACT SITE 204 204 PROTON DONOR (PROBABLE).
FT MUTAGEN 200 200 D->N: DECREASE IN ACTIVITY.
FT MUTAGEN 200 200 D->E: NO CHANGE IN ACTIVITY.
FT MUTAGEN 204 204 E->D,Q: LOSS OF ACTIVITY.
SQ SEQUENCE 699 AA; 73677 MW; AC7C9B2E2987643 CRC64;

Query Match 19.4%; Score 484; DB 1; Length 699;
Best Local Similarity 27.4%; Pred. No. 3.2e-27;
Matches 130; Conservative 80; Mismatches 156; Indels 108; Gaps 18;

QY 1 YNLICYFTNWAQYRGLSGFKPDINPCLCTHLYAFAGM----- 40
DB 44 YKIVGYPSWAAG--RNVNVAIDIDPTKVTHINAFADICWNGHGNPDSPGNPVTWT 100

QY 41 -ONNEITIE-----WNDVLYKAF-----NDLKNRNSKLTL 72
DB 101 CONEKSQTINVPNGTIVLGDPMIDTG--KTFAGDTWQDFIAGNINQLNKLKQTNPNLKI 158

QY 73 LAIGGNFETAPPTMTWTSQNRQTFTSVTKFLROYGFDGLDLDWEYVPGSG 127
DB 159 ISVGGWTWSNR-FSDVAATAATREVFANSADVFLKYNFDFGLDWEYFVSGGLDGNKR 217

QY 128 PQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELSKYLDPIH 187
DB 218 PEDKQNYTLTLLSKIREKLDAAAGVDGKKYLLTIA--SCASATYAANTELAIAIVDWIN 275

QY 188 VMTYDLHGSWEGYTGNSPLYKYPTET-----GSNAYLVNDVVMYKNNNG 233
DB 276 IMTYDFNGAWOKISAHNAFLNYPDAASAAGVDPDANTFNVAAGAQHLLDA-----G 325

QY 234 APAEKLIVGFPEYGHFTILRNPDSNGIGATSGDPAGATROAGFWAYEICTFLR 291
DB 326 VPAKLVGLVFPYGRGWGCGCAAGNGQYQTCGTGSSVG--TWEAGSFDYDLEANYINKN 383
```



Search completed: June 29, 2003, 21:00:16  
Job time : 10.0237 secs

## RESULT 15

RP SEQUENCE OF 35-67 FROM N.A., AND SEQUENCE OF 35-67.

SEQUENCE 1023479; R001347; PubMed=1735126;  
MEDLINE=92136357; R001347; PubMed=1735126;  
RX Blaiseau P.-L., Kunz C., Grison R., Bertheau Y., Brygoo Y.;  
RA Blaiseau P.-L., Kunz C., Grison R., Bertheau Y.;  
RT "Cloning and expression of a chitinase gene from the hyperparasitic  
fungus *Aphanocladium album*";  
RL Curr. Genet. 21:61-66(1992).  
CC -|- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
CC acetyl-D-glucosamine polymers of chitin.  
CC -|- SUBCELLULAR LOCATION: Secreted.  
CC -|- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----

CC	EMBL; X64104; CAA45468.1; -	
DR	EMBL; S81303; AAB21333.1; -	
DR	PIR; JQ1975; JQ1975.	
DR	HSSP; P07254; 1CTN.	
DR	InterPro; IPR001579; Chitinase 18/2.	
DR	InterPro; IPR001223; Glyco_hydro_18.	
DR	Pfam; PF00704; Glyco_hydro_18; 1.	
DR	ProDom; PD000471; Glyco_hydro_18; 1.	
DR	PROSITE; PS01095; CHITINASE 18; 1.	
KW	Hydrolase; Glycosidase; Chitin degradation; Signal; Zymogen.	
FT	SIGNAL 1 22	POTENTIAL.
FT	PROPEP 23 34	POTENTIAL.
FT	CHAIN 35 423	CHITINASE 1.
FT	ACT SITE 171 171	PROTON DONOR (BY SIMILARITY).
SO	SEQUENCE 423 AA; 46058 MW; 55DFB25B73443F31	CRC64:

Query Match  
18.7%; Score 466.5; DB 1; Length 423;

Best Local Similarity 30.0%; Pred. No. 3e-26; Matches 115; Conservative 64; Mismatches 157; Indels 47; Gaps 14;

Qy 6 YFTNWAQYRPGLSGFKPDDINPCLCTHLIYAFAGMQ-----NNEITTEWN 51  
||||| :||| :||| :||| :|||

**THIS PAGE BLANK (08070)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:10:28 ; Search time 1489.73 Seconds  
(without alignments)  
4913.891 Million cell updates/sec

Title: US-10-004-219b-9  
Perfect score: 2493  
Sequence: 1 YNLICYFTNWAQYRFLGSLF.....QQHCQAGLVFTSCNCCNWP 452

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame.p2n.model -DEV=xlp  
-Q/cg2\_1/USPTO\_spool/US10004219/runat\_24062003\_160228\_8173/app.query.fasta\_1.1877  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CGN 1.1 4575 @runat\_24062003\_160228\_8173 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*\*  
1: em\_estba:\*\*  
2: em\_esthum:\*\*  
3: em\_estin:\*\*  
4: em\_estmu:\*\*  
5: em\_estov:\*\*  
6: em\_estpl:\*\*  
7: em\_estro:\*\*  
8: em\_hctc:\*\*  
9: gb\_est1:\*\*  
10: gb\_est2:\*\*  
11: gb\_htc:\*\*  
12: gb\_est3:\*\*  
13: gb\_est4:\*\*  
14: gb\_est5:\*\*  
15: em\_estfun:\*\*  
16: em\_estom:\*\*  
17: gb\_gss:\*\*  
18: em\_gss\_hum:\*\*  
19: em\_gss\_inv:\*\*  
20: em\_gss\_pln:\*\*  
21: em\_gss\_vrt:\*\*  
22: em\_gss\_fun:\*\*  
23: em\_gss\_mam:\*\*  
24: em\_gss\_mus:\*\*  
25: em\_gss\_othr:\*\*  
26: em\_gss\_pro:\*\*  
27: em\_gss\_rod:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2488	99.8	1533	11	AK008633	Mus muscu
2	2467	99.0	1535	11	AK008757	Mus muscu
3	2416	96.9	1534	11	AK008650	Mus muscu
4	2365.5	94.9	1530	11	AK007573	Mus muscu
5	1498	60.1	892	14	BQ231098	AGENCOURT
6	1452	58.2	900	12	BG867646	BG867646 602787435
7	1431	57.4	838	12	BG867815	BG867815 602786336
8	1416	56.8	809	12	BG866417	BG866417 602785352
9	1414	56.7	843	12	BG868207	BG868207 602784727
10	1400	56.2	872	14	BQ231064	AGENCOURT
11	1395	56.0	923	13	B1686954	B1686954 603313831
12	1394.5	55.9	830	12	BG865815	BG865815 602788220
13	1381.5	55.4	1007	9	AV072529	AV072529 AV072529
14	1363	54.7	857	12	BG871434	BG871434 602790636
15	1353	54.3	800	12	BG866853	BG866853 602786589
16	1349	54.1	763	12	BG872647	BG872647 602793705
17	1348	54.1	939	14	BQ219632	AGENCOURT
18	1346	54.0	764	12	BG869304	BG869304 602788819
19	1346	54.0	837	12	BG868741	BG868741 602787570
20	1341	53.8	894	12	BG869558	BG869558 602789826
21	1339	53.7	855	12	BG872136	BG872136 602792374
22	1335.5	53.6	949	12	BG868019	BG868019 602788381
23	1333	53.5	1100	12	BF302298	BF302298 602031269
24	1330.5	53.4	886	12	BG870840	BG870840 602791891
25	1330	53.3	808	12	BF534404	BF534404 602047909
26	1330	53.3	844	12	BF536293	BF536293 602051942
27	1324	53.1	757	12	BG867292	BG867292 602788635
28	1321.5	53.0	880	12	BG871468	BG871468 602790678
29	1321.5	53.0	1155	12	BG867473	BG867473 6027889716
30	1315.5	52.8	836	12	BG869655	BG869655 602789716
31	1315.5	52.8	848	12	BG869143	BG869143 602789020
32	1313.5	52.7	874	12	BG865756	BG865756 602784345
33	1311	52.6	833	12	BG871337	BG871337 602790509
34	1309	52.5	830	12	BG869896	BG869896 602789435
35	1308.5	52.5	882	12	BG873488	BG873488 602791730
36	1306.5	52.4	788	12	BG873335	BG873335 602794304
37	1305	52.3	847	12	BG866775	BG866775 602786605
38	1305	52.3	856	12	BG870147	BG870147 602793390
39	1302	52.2	948	13	B1664452	B1664452 603290028
40	1299.5	52.1	845	12	BG865848	BG865848 602788264
41	1299	52.1	976	14	BQ225554	AGENCOURT
42	1297.5	52.0	863	12	BG872850	BG872850 602793973
43	1296	52.0	742	12	BG873143	BG873143 602794547
44	1293.5	51.9	878	12	BG868262	BG868262 602784787
45	1293	51.9	828	12	BG869811	BG869811 602789327

ALIGNMENTS

RESULT 1  
AK008633  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

AK008633  
Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:chitinase, acidic, full insert sequence.  
AK008633.1  
GI:12842941  
HTC; CAP trapper.  
Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library  
clone:2200003E03.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1533 bp  
mRNA  
linear  
HTC 19-JAN-2002

**AUTHORS**  
**TITLE**  
**JOURNAL**  
**MEDLINE**  
**PUBMED**

Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

REFERENCE	JOURNAL
AUTHORS	MEDLINE
TITLE	PUBMED
	REFERENCE
	AUTHORS

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
20459374  
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.

REFERENCE  
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuaki, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hayada, A.,  
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
1. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
2. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
3. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
4. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
5. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
6. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
7. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
8. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
9. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...
10. The effect of the ...	Journal of ...	12345678	12345678	12345678	Smith J, ...

RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsumoto, Y., Nikaudo, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hoernann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Mazzionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Satoh, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Totoy-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,  
Wysshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsuke, S.,  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851  
 5 (bases 1 to 1533)  
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
 Arakawa, T., Baldarelli, P., Bono, H., Brownstein, M., Bult, C.,  
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagata, T.,  
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,  
 Hume, D., Imotani, K., Ishii, Y., Itomoto, K., Izawa, M., Kasukawa, T.,  
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
 Kurihara, C., Matsuayama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
 Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
 Saito H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
 Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,  
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
 Hayashizaki, Y.

Db 260 GACTTTGAAAAACAGGAACAGCAAACTGAAACCCCTCTGGCAATGGAGGCTTGAACCTTT 319  
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100  
Db 320 GGAATCTCTCTTTTCACTACCACTGGTTTCCACTTCTCAGAACCCGACACCTTCATTACC 379  
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 380 TCAGTCATCAAAATTTCTGGCTCAGTATGGTTTGTATGGACTGGAGCTGGGAATAC 439  
QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140  
Db 440 CCAGGCTCACGTGGGAGCCCTCTCAGGACAAAGCATCTTCACTGTCTCGTGAAGGAA 499  
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160  
Db 500 ATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTTACT 559  
QY 161 AlaAlaValAlaGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
Db 560 GCTGCTGTAGCTGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCTCGAATTTCT 619  
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
Db 620 AAGTACCTGGATTTCACTCATGTCATGACATGATGACCTCCATGGCTCTCTGGGAGGCTAC 679  
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220  
Db 680 ACTGGGGAGAAAGTGCCTCTTTACAAATACCTTACTGAGACTGGTAGCAATGCCATCTTC 739  
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnGlnValAlaProAlaGluLysLeuIle 240  
Db 740 ATGTGGATTATGTCATGAACTATTGGAAGACAACTGGAGCCCGAGCTGAGAAGCTCAT 799  
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260  
Db 800 GTTGGATTTCCAGAGTATGGACACACTTTCATCTCCTGAGAAACCCCTCTGTAATGGAAT 859  
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280  
Db 860 GTGCCCCCTACTCTGTGTGATGCCCTCTGTGCCCTATACACAGACAGCTGGGTCTGG 919  
QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300  
Db 920 GCTACTATGAGATTGCACCTTCTGAGAGTGGAGCCACTGAGGTCTGGATGCCTCC 979  
QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320  
Db 980 CAAGAAGTGCCTATGCTTATAAGGCCAACGAGTGGCTTGGCTATGACAAATATCAAGAGC 1039  
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340  
Db 1040 TTCAGTGTAAAGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCC 1099  
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360  
Db 1100 ATGACCTTGATGACTTCACTGGCTTTTCTGTGATCAGGAAATTTCTCTGACTTCT 1159  
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380  
Db 1160 ACTTTGAACAAAGCCCTTGGCATATCACTGAAGGTGTGCACAGCTCTCGACGTGCCTTCC 1219  
QY 381 GluProValThrThrProGlySerGlySerGlyGlyGlySerSerGlySerSer 400  
Db 1220 GAGCCAGTACTACTCTCCAGAAAGTGGAGTGGGGGTGGAGGTCCCGAGGAAGCTCT 1279  
QY 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420  
Db 1280 GGAGGCGATGGATTTGTGCCCAAGCAGATGGCTCTTACCCTGTGCAGATGACAGA 1339  
QY 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440

Db 1340 AATGCTTTTGGCAGTCATCAATGGAATCATACACCAGCAGCATTTGTCAAGAGGGCTT 1399  
QY 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452  
Db 1400 GTTTTGTATACAGCTGTAAATGCTGCAACTGGCCA 1435  
RESULT 2  
LOCUS AK008757  
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210020N23;chitinase, acidic, full insert sequence.  
ACCESSION AK008757  
VERSION AK008757.1 GI:12843148  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:CS7BL/6J) adult male stomach cDNA to mRNA, clone:lib-RIKEN full-length enriched mouse cDNA library  
clone:2210020N23.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
3  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
4  
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C., Fletcher, M., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409 (6821), 685-690 (2001)  
MEDLINE 21085660  
PUBMED 11217851  
5 (bases 1 to 1535)  
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,











```
Db 915 GCCTACTATGAGATTGACCTTTCTGAGAAGTGGAGCCACTGAGCTCTGGGATGCTCC 974
Qy 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
Db 975 CAAGAAGTGGCCCTATGCTATAGGCCAAGCAGTGGCTTGGCTATGACAAATATCAAGAGC 1034
Qy 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
Db 1035 TTCAGTGTAAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCC 1094
Qy 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
Db 1095 ATTGACCTTGATCACTTCTGCTCTTCTGTGATCAGGAAATTTCTCTGACTTCT 1154
Qy 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
Db 1155 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCTGACGGCTTCC 1214
Qy 381 GluProValThrThrProProGlySerGlySerGlyGlyGlySerSerGlyGlySerSer 400
Db 1215 GAGCCAGTGACTACTCTCCAGAAAGTGGAGTGGGGTGGAGCTCCGAGGAGACTCT 1274
Qy 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420
Db 1275 GGAGGCAGTGGATTCTGTGCGCAGAAAGCAGATGGCTCTTACCCTGTGGCAGATGACAGA 1334
Qy 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440
Db 1335 AATGCTTTTGGCAGTGATCAATGAATACATACAGCAGCATGTCAAGCAGGGCTT 1394
Qy 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
Db 1395 GTTTTGTATACCACTGTAATTTCTGCACTGGCCA 1430

RESULT 5
BQ231098 892 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT 7578138 NCI CGAP St1 Mus musculus cDNA clone
DEFINITION IMAGE:6051580 5', mRNA sequence.
ACCESSION BQ231098
VERSION BQ231098.1 GI:20412498
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 892)
NIH-MGC http://mgi.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13305 row: k column: 05
High quality sequence stop: 772.
Location/Qualifiers
1..892
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6051580"
/clone_lib="NCI CGAP St1"
/lab_host="DH10B (Tr-resistant)"
/note="Organ: stomach; Vector: pCMV-Sport6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.77 kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
```

## FEATURES

Source

BASE COUNT 219 a 226 c 216 g 230 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 892  
Score: 1498.00 Matches: 279  
Percent Similarity: 98.94% Conservative: 0  
Best Local Similarity: 98.94% Mismatches: 1  
Query Match: 60.09% Indels: 2  
DB: 14 Gaps: 0

US-10-004-219B-9 (1-452) x BQ231098 (1-892)

Qy 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20  
Db 36 TACATCTGATGATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGCTCGGGAGCTTC 95  
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 96 AAGCCTGATGATCAATTAACCCCTGCTGTGATCTACCTGATCTATGCTCTTCTGGGATG 155  
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60  
Db 156 CAGAACATGATGATCACCACCATAGATGAATGAATGATTTACTCTCTATAAAGCTTTCAAT 215  
Qy 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
Db 216 GACTTGAAAACAGAGAAACAGCAACTGAAACCCCTCTGCAATTTGGAGGCTGGAACTTT 275  
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100  
Db 276 GGAACCTGCTCTCTTTCACCTACCATGCTTTCCACTCTCAGAACCCGACAGCTTCATTACC 335  
Qy 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 336 TCAGTCATCAAAATTTCTGCGTCAGTATGGGTTGATGAGCTGGAGCTGGAGCTGGGAAATAC 395  
Qy 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140  
Db 396 CCAGGCTCAGCTGGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGCTCTGGTGAAGAA 455  
Qy 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160  
Db 456 ATGCGTGAAGCTTTTGGCAGGAGGCTATTGAGACCAACAGGCCAGACTGATGCTTACT 515  
Qy 161 AlaAlaValAlaGlyGlyLysSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
Db 516 GCTGCTGTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCT 575  
Qy 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
Db 576 AAGTACCTGGATTTTCATCATGTCATGACATATGACCTCCATGGCTCTGGAGGGCTTAC 635  
Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220  
Db 636 ACTGGGAGAGATAGTCTCTTTACAAATACCTACTGAGACTGGTAGCAATGCCCTACCTC 695  
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAsnGlnGlyAlaProAlaGluLysLeuIle 240  
Db 696 AATGTGGATTATGTCATGAACCTATTGGGAAGAACAAATGGAGCCCGAGCTGAGAGCTCA 755  
Qy 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260  
Db 756 GTTGGATTTCCAGAGTATGGACACACCTTCATCTCGAAGAACCCCTCTGATATATGAAT 815  
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGln-AlaGly-PheThr 280  
Db 816 GGTGCCCCCTACTCTGTGTGATGGCCCTCTGCTGGCCCTATACAGACAGGCTGGNGTTCT 875  
Qy 280 tp 280  
Db 876 GG 877

```

RESULT 6
LOCUS      BG867646
DEFINITION 602787435F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4913288 5',
            linear mRNA EST 29-MAY-2001
ACCESSION  BG867646
VERSION     BG867646.1 GI:14218186
KEYWORDS   EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 900)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10818 row: f column: 09
            High quality sequence stop: 852.
FEATURES    Location/Qualifiers
             1..900
             /organism="Mus musculus"
             /strain="FVB/N"
             /db_xref="taxon:10090"
             /clone="IMAGE:4913288"
             /clone_lib="NCI_CGAP_SG2"
             /lab_host="DH10B (T1 phage-resistant)"
             /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
             Note1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo
             dt. Average insert size 1.3 kb. Constructed by Life
             Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT  218 a 217 c 234 g 231 t
ORIGIN
Alignment Scores:
Pred. No.:      1,27e-143      Length:      900
Score:          1452.00      Matches:    289
Percent Similarity: 95.71%      Conservative: 1
Best Local Similarity: 95.38%      Mismatches: 3
Query Match:    58.24%      Indels:    10
DB:             12          Gaps:      0
US-10-004-219B-9 (1-452) x BG867646 (1-900)
QY 148 GluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIle 167
Db 1 GAGGCTATTGAGAGCAACAGCCGACAGCTGATGGTTACTGCTGCTGCTAGCTGGTGGATT 60
QY 168 SerAsnIleGluAlaGlyTyrGluIleProGluLeuSerIleValThrLeuAspPheIleHis 187
Db 61 TCCAAATCCAGCTGGCTGCTGATGAGATCCCTGACATTTCTTAAGTACTGATGATTCATCCAT 120
QY 188 ValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyLeuAsnSerProLeu 207
Db 121 GTCATGACATATGACTCCATGCTCTGCTGGAGGCTACACTGGGAGAGATAGTCTCTT 180
QY 208 TyrIleValProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsn 227
Db 181 TACAAATACCTTACTGAGACTGCTGATGATGCTTACCTCAATGTGATTTATGTCATGAAC 240
QY 228 TyrTrpIleAsnAsnGlyAlaProAlaGluIleValGlyPheProGluTyrGly 247
Db 241 TATTGGAAGAAACAATGGAGCCCCAGCTGAGAGCTCATTTGTTGGATTCACGAGATGGA 300
QY 248 HisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAsp 267

```

```

Db 301 CACACCTTCATCTCTGAGAAACCCCTCTGATAATGGAATGGTGGCCCTACCTCTGGTGTAT 360
QY 268 GlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheThrAlaTyrTyrGluIleCysThr 287
Db 361 GGCCCTGCTGGGCCCTATACCAGACAGGCTGGTCTTGGGCTACTATGAGATTGGACCC 420
QY 288 PheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyrAlaTyr 307
Db 421 TTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCCCAAGAAGTGGCCCTATGCCTAT 480
QY 308 LysAlaAsnGluTrpLeuGlyTyrAspAsnIleIysSerPheSerValIleAsnGlnTrp 327
Db 481 AAGGCCAACGAGTGGCTTGGCTATGACATATCAAGAGCTTCAGTGTTAAGCTCAGTGG 540
QY 328 LeuIysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspAspPheThr 347
Db 541 CTTAAGCAGAAACAATTTGGAGGTGCCATGATCTGGGCCATGACCTTGATGACTTCACT 600
QY 348 -GlySerPheCysAspGlnGly-LysPheProLeuThrSerThr-LeuAsnLysAlaLeu 366
Db 601 AGGCTCTTTCTGTGATCAGGGACAAATTTCTCTGACTTCTACTTTTGAACAAGCCCTT 660
QY 367 GlyIleSerThrGlu-GlyCysThrAlaProAspValProSerGluProValThrThrPr 386
Db 661 GGCATATCCACTGAACGGTTGCACAGCTCCTGACGTGCTTCCGAGCCAGTGACTACTCC 720
QY 386 oProGlySerGlySerGlyGlySerSer-GlyGlySerSerGlyGlySerGlyPheC 406
Db 721 TCCAGGAAGTGGGAATGGGGTGGGAAGCTCCGGGAGGAAGCTCTGGAGGAGTGATTTCT 780
QY 406 ysAlaAspLysAla-AspGlyLeuTyrPro-ValAlaAspAspArg-AsnAlaPheTrpG 425
Db 781 GTGCGCAGAACGACAGATGGCTCTTAACCTTGTGCAGATGACCGAANAATGCTTTTGGC 840
QY 425 InCysIle-AsnGlyIle-ThrTyrGlnGlnHisCysGlnAlaGlyLeu 440
Db 841 AGTGATTCATTAATGGAATTCACATACAGCAGCATTTGTCCAGCAGGCTTG 889
RESULT 7
BG867815
LOCUS      602786336F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912544 5',
            linear mRNA sequence.
ACCESSION  BG867815
VERSION     BG867815.1 GI:14218355
KEYWORDS   EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 838)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10816 row: g column: 09
            High quality sequence stop: 824.
FEATURES    Location/Qualifiers
             1..838
             /organism="Mus musculus"
             /strain="FVB/N"
             /db_xref="taxon:10090"
             /clone="IMAGE:4912544"
             /clone_lib="NCI_CGAP_SG2"
source

```

/lab\_host="DH10B (T1 phage-resistant)"  
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:  
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT: 209 a 211 c 198 g 220 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1 1.94e-141 Length: 838  
Score: 1431.00 Matches: 263  
Percent Similarity: 98.88% Conservativeness: 2  
Best Local Similarity: 98.13% Mismatches: 3  
Query Match: 57.40% Indels: 0  
DB: 12 Gaps: 0

US-10-004-219B-9 (1-452) x BG867815 (1-838)

```
QY 1 TyrAsnLeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
DB 34 TACAATCTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGGAGCTTC 93
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMet 40
DB 94 AAGCCTGATGACATTAAACCCCTGCTGTGTACTCCTGATCTATGCTTGTCTGGGATG 153
QY 41 GlnAsnAsnGluileThrThrileGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
DB 154 CAGAACATGAGATCACCACCATAGATGGAATGATGTTACTCTCTATAAAGCTTCAAT 213
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyGlyTrpAsnPhe 80
DB 214 GACTTGAAAAACAGGAACAGCAAACTGAAACCCCTCTGGCAATTGGAGGCTGGAACCTT 273
QY 81 GlyThrAlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheileThr 100
DB 274 GGAACCTGCTCTCTTCACTACCATGGTGTTCACATCTCTCAGAACCCGAGACCTTCATTAC 333
QY 101 SerValileLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 334 TCAGTCATCAAAATTTCTGGCTCAGTATGGTTGATGGACTGGACCTGGAGCTGGGATAC 393
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
DB 394 CCAGGCTCAGTGGGAGCCCTCTCAGSACAAGCATCTCTCACTGCTCTGGTGAAGGAA 453
QY 141 MetArgGluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThr 160
DB 454 ATCGGTGAAGCTTTTGAGCAGGAGGGCTATTGAGAGCAACAGGCCCAAGACTGATGGTTACT 513
QY 161 AlaAlaValAlaGlyGlyLysSerAsnileGlnAlaGlyTyrGluileProGluLeuSer 180
DB 514 GCTGCTGTAGCTGTGGATTTTCCACATCCAGGCTGGCTATGAGATCCTGNACTTCT 573
QY 181 LysTyrLeuAspPheileHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
DB 574 AAGTACCTGGATTTTCAATGTCATGACATATGACCTCCATGGCTCTCTGGAGGGCTAC 633
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
DB 634 ACTGGGGAGAAATAGTCTCTTTTACAATACCCCTACTGAGACTGGTAGCAATGCTACCTC 693
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyValaProAlaGluLysLeuile 240
DB 694 AATGTGATTAATGTCATGAATGATGGAAGAACATGAGGCCCACTGAGAGCTCAAT 753
QY 241 ValGlyPheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyIle 260
DB 754 GTTGGATTCCAGAGTTTGACACACCTTCTATCCTGAGAAACCCCTCTGATATGATTT 813
QY 261 GlyAlaProThrSerGlyAspGly 268
DB 814 GGTGCCCTTACTCTGTGTGAATGTC 837
```

RESULT 8  
BG866417  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

US-10-004-219B-9 (1-452) x BG866417 (1-809)  
Pred. No.: 7 21e-140 Length: 809  
Score: 1416.00 Matches: 264  
Percent Similarity: 98.14% Conservativeness: 0  
Best Local Similarity: 98.14% Mismatches: 4  
Query Match: 56.80% Indels: 1  
DB: 12 Gaps: 0

BASE COUNT  
ORIGIN

US-10-004-219B-9 (1-452) x BG866417 (1-809)  
QY 71 ThrLeuLeuAlaileGlyGlyTrpAsnPheGlyThrAlaProPheThrThrMetValSer 90  
DB 2 ACCCTCTGGCAATTTGGAGGCTGGAACCTTGAACCTGCTCTTCACTACCATGGTTCC 61

QY 91 ThrSerGlnAsnArgGlnThrPheileThrSerValileLysPheLeuArgGlnTyrGly 110  
DB 62 ACTTCTCAGAACCGCCAGACCTTCACTACCTCAGTCATCAAAATTTCTGCTCAGTATGG 121  
QY 111 PheAspGlyLeuAspLeuAspTrpGluTyrProGlySerArgGlySerProGlnAsp 130  
DB 122 TTTGATGACTGGACCTGGACTGGCAATACCCAGGCTCAGCTGGAGGCTCTCTCAGGAC 181  
QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaile 150  
DB 182 AAGCATCTCTTCACTGCTCTGCTGGAAGAAATGCGTGAAGCTTTTTCAGCAGGAGCTATT 241  
QY 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsnile 170  
DB 242 GAGAGCAACAGGCCCAAGCTGATGTTACTGCTGTAGCTGGTGGGATTTTCCACATC 301

BG866417 809 bp mRNA linear EST 29-MAY-2001  
602785352F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4911421 5',  
mRNA sequence.

BG866417 GI:14216957  
BG866417

EST.  
house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgs.nci.nih.gov/.  
1 (bases 1 to 809)

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
Plate: LLAM10813 row: h column: 14  
High quality sequence stop: 804.

Location/Qualifiers  
1. 809

/organism="Mus musculus"  
/strain="FVB/N"

/db\_xref="taxon:10090"  
/clone="IMAGE:4911421"

/clone\_lib="NCI CGAP SG2"  
/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:  
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

193 a 200 c 208 g 207 t  
I others

Qy 171 GlnAlaGlyTyrGluLeuProGluLeuSerTyrLeuAspPheHisValMetThr 190  
 Db 302 CAGGCTGGCTATGAGATCCCTGAACCTTCTAAGTACCTGGATTTCATCATGTCATGACA 361  
 Qy 191 TyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrIlyThr 210  
 Db 362 TATGACCTCCATGGCTCTCTGGAGGGCTACACTGGGAGAAATAGTCTCTTTACAAATAC 421  
 Qy 211 ProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrIlyThr 230  
 Db 422 CCTACTGAGACTGGTAGCAATGCCCTACCTCATGTGGATTATGATCATGAACTATTGGAG 481  
 Qy 231 AsnAsnGlyAlaProAlaGluLysLeuLeuValGlyPheProGluTyrGlyHisThrPhe 250  
 Db 482 AACATGGAGGCCAGCTGAGAAGCTCATTTGTGGATTCCCAAGATGAGACACACCTTC 541  
 Qy 251 IleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyProAla 270  
 Db 542 ATCCCTGAGAAACCCCTCTGATTAATGGAATGGTGGCCCTACCTCTGGTGGTGGCCCTGCT 601  
 Qy 271 GlyAlaTyrThrArgGlnAlaGlyPheTyrPalaTyrTyrGluLeuLeuCysThrPheLeuArg 290  
 Db 602 GGGCCCTATACAGACAGCTGGTCTGGGCTACTATGATGATGACCTTTCTGAGA 661  
 Qy 291 SerGlyAlaThrGluValTyrAspAlaSerGlnGluValProTyrAlaTyrLysAlaAs 310  
 Db 662 AGTGAGCCACTGAGGCTCTGGGATGCCCTCCCAAGAGTGGCCCTATGCTATAAGGCCAA 721  
 Qy 310 nGluTyrLeuGlyTyrAspAsnIleGlySerPheSerValLysAlaGlnTyrLeuLysG 330  
 Db 722 CGAGTGGCTGGCTATGATCATATCAAGAGCTTCAGTGTAAAGGTCAGTGGCTTTAAGCA 781  
 Qy 330 nAsnAsnPheGlyGlyAlaMetIle 338  
 Db 782 GAACAATTTGGAGTGCCATGATC 806

RESULT 9  
 LOCUS BG868207  
 DEFINITION 602784727F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4910958 5',  
 mRNA sequence.  
 ACCESSION BG868207  
 VERSION BG868207.1 GI:14218747  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM10812 row: e column: 07  
 High quality sequence stop: 792.  
 Location/Qualifiers  
 1..843  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4910958"  
 /clone\_lib="NCI CGAP\_SG2"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 204 a 207 c 217 g 215 t

ORIGIN

Alignment Scores:  
 Pred. No.: 1.26e-139 Length: 843  
 Score: 1414.00 Matches: 275  
 Percent Similarity: 96.83% Conservative: 0  
 Best Local Similarity: 96.83% Mismatches: 4  
 Query Match: 56.72% Indels: 6  
 Gaps: 0

US-10-004-219B-9 (1-452) x BG868207 (1-843)

Qy 111 PheAspGlyLeuAspLeuAspTyrGluTyrProGlySerArgGlySerProProGlnAsp 130  
 Db 2 TTTGATGGACTGGAGCTGGACTGGGAATACCAGGCTCACGTGGAGCCCTCTCAGGAC 61  
 Qy 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150  
 Db 62 AAGCATCTCTTCACTGTCTGTGTGAAGAAATGCGTGAAGCTTTTGACAGAGGCTATT 121  
 Qy 151 GluSerAsnArgProArgLeuMetValThrAlaValAlaGlyGlyIleSerAsnIle 170  
 Db 122 GAGAGCAACAGGCCAGACTGATGCTGTCTGTAGCTGGTGGGATTTTCCAACATC 181  
 Qy 171 GlnAlaGlyTyrGluLeuProGluLeuSerIlyThrLeuAspPheIleHisValMetThr 190  
 Db 182 CAGGCTGGCTATGAGATCCCTGAACTTTCTAAGTACCTGGATTTTCATCATGTGATGACA 241  
 Qy 191 TyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrIlyThr 210  
 Db 242 TATGACCTCCATGGCTCTCTGGAGGGCTACACTGGGAGATAGTCTCTTTTACAAATAC 301  
 Qy 211 ProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrIlyThr 230  
 Db 302 CCTACTGAGACTGGTAGCAATGCCCTACCTCAATGGATTATGTCATGAACATATTGGAG 361  
 Qy 231 AsnAsnGlyAlaProAlaGluLysLeuLeuValGlyPheProGluTyrGlyHisThrPhe 250  
 Db 362 AACATGGAGGCCAGCTGAGAAGCTCATTTGTGGATTCCCAAGATGAGACACACCTTC 421  
 Qy 251 IleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyProAla 270  
 Db 422 ATCCCTGAGAAACCCCTCTGATTAATGGAATGGTGGCCCTACCTCTGGTGGTGGCCCTGCT 481  
 Qy 271 GlyAlaTyrThrArgGlnAlaGlyPheTyrPalaTyrTyrGluLeuLeuCysThrPheLeuArg 290  
 Db 482 GGGCCCTATACAGACAGCTGGTCTGGGCTACTATGAGATTTCACACCTTTCTGAGA 541  
 Qy 291 SerGlyAlaThrGluValTyrAspAlaSerGlnGluValProTyrAlaTyrLysAlaAs 310  
 Db 542 AGTGAGCCACTGAGGCTGGAGTGGCCCTCCCAAGAGTGGCCCTATGCTATAAGGCCAA 601  
 Qy 311 GluTyrLeuGlyTyrAspAsnIleLysSerPheSerValLysAlaGlnTyrLeuLysGln 330  
 Db 602 CAGTGGCTGGCTATGACAATATCAAGAGCTTCAGTGTAAAGGCTCAGTGGCTTAAGCAG 661  
 Qy 331 AsnAsnPheGlyGlyAlaMetIleTyrPalaIleAspLeuAspAspPhe-ThrGlySerPh 350  
 Db 662 AACAAATTTGGAGGTGCCATGATCTGGGCCATTGACCTTGTATGACTTCAACTGGGCTCTTT 721  
 Qy 350 eCysAspGlnGly-LysPhePro-LeuThrSer-ThrLeuAsn-LysAlaLeuGlyIle 369  
 Db 722 CTGTGATCAGGGAATAATTTCTTCTGACTTCTAACTTTGAACAAAGCCCTGGGCAATAT 781  
 Qy 369 eThrGluGlyCysThrAlaProAspValProSerGluProValThrThrProGlyLys 389  
 Db 782 CCAGTGGAGTTGGACAGT-CCTGACGTGGCTTCGAGCCAGTGGTACTACTCTCTCCAGGAA 840  
 Qy 389 er 389  
 ||

```

Db      841 GT 842
RESULT 10
BQ231064
DEFINITION: BQ231064 872 bp mRNA linear EST 02-MAY-2002
IMAGE:6051490 5', mRNA sequence.
ACCESSION BQ231064
VERSION BQ231064.1 GI:20412464
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 872)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13305 row: g column: 11
High quality sequence stop: 724.
Location/Qualifiers
1. .872
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6051490"
/lab_host="NCI CGAP St1"
/notes="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.77 kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 218 a 222 c 205 g 225 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 4,09e-138 Length: 872
Score: 1400.00 Matches: 262
Percent Similarity: 98.87% Conservative: 0
Best Local Similarity: 98.87% Mismatches: 1
Query Match: 56.16% Indels: 2
DB: 14 Gaps: 0
US-10-004-219b-9 (1-452) x BQ231064 (1-872)
QY 1 TyrAsnLeuLeuCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 68 TACAATCTGATATGCTATTTACCAACTGGGCCAGTATCGGCAGGCTCTGGGAGCTTC 127
QY 21 LysProAspAspLeuProCysLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40
Db 128 AAGCCTGATGACATTAACCCCTGCTGTACTCCTGATCTATCCCTTTGCTGGGATG 187
QY 41 GlnAsnAsnGluLeuThrThrLeuGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 188 CAGAACATGAGATCACCACCAGTAAATGATGATGTTACTCTCTATTAAGCTTTCAT 247
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaLeuGlyTrpAsnPhe 80
Db 248 GACTTGAAAAACAGGACACGAACTGAAACCCCTCTGGCAATTTGGAGGCTGGAATTT 307
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheThr 100
Db 308 GGAACTGCTCTTTTCACTACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATTACC 367

```

101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120  
 368 TCAGTCATCAATTTCTGGTCAGTATGGTTTGTAGTGGACCTGGACCTGGGAATAC 427

121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140  
 428 CCAGGCTCACGTGGGAGCCCTCCTCAGGACCAAGCATCTCTTCACTGTCTCTGGTGAAGGAA 487

141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160  
 488 ATGCGTGAAGCTTTTGACGAGGAGGCTATTGAGACCAACAGGCCCCAGACTGATGTTACT 547

161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180  
 548 GTCTGTGTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCT 607

181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
 608 AAGTACCTGGATTTTCATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 667

201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220  
 668 ACTGGGAGATGATGCTCTTTTACAAATACCTACTGAGACTGGTAGCAATGCCCTACCTC 727

221 AsnValAspTyrValMetAsnTyrTrpLysAsnGlyAlaProAlaGluLysLeuIle 240  
 728 AATGTGGATTATGTCATGAACCTATTGGAAGAACAAATGAGCCCAAGCTGAGAAGCTCAT 787

241 ValGlyPheProGluTyrGlyHisThr-PheIleLeuArgAsnProSerAspAsn-GlyI 260  
 788 GTTGGATTCCAGAGTATGACACACCCCTTCATCTGAGAAACCCCTCTGTAATATGGGAA 847

260 LeGlyAlaPro 263  
 848 TTGGTGGCCCCC 858

RESULT 11  
 B1686954  
 LOCUS B1686954  
 DEFINITION 603313831F1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:5353927 5',  
 mRNA sequence.  
 ACCESSION B1686954  
 VERSION B1686954.1 GI:15649582  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Jeffrey Green M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM11899 row: n column: 08  
 High quality sequence stop: 795.  
 Location/Qualifiers  
 1. .923  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5353927"  
 /clone\_lib="NCI CGAP Mam6"  
 /sex="female, virgin"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"

FEATURES  
 source

/lab host="DH10B"

/note="organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
Library constructed by Life Technologies. Investigator  
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 221 a 232 c 241 g 229 t

ORIGIN

## Alignment Scores:

Pred. No.: 1 52e-137 Length: 923  
Score: 1395.00 Matches: 277  
Percent Similarity: 90.00% Conservativity: 2  
Best Local Similarity: 89.35% Mismatches: 21  
Query Match: 55.96% Indels: 11  
DB: 13 Gaps: 2

US-10-004-219B-9 (1-452) x B1686954 (1-923)

QY 126 SerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGluAlaPhe 145  
DB 2 AGCCCTCTCAGCAAGCATCTCTTCACTGCTGCTGGAAGAAATGCTGAAGCTTTT 61  
QY 146 GluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGly 165  
DB 62 GAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTTACTGCTGCTAGCTGGT 121  
QY 166 GlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPhe 185  
DB 122 GGGATTTCAACATCCAGCTGGCTATGAGATCCCTGAACTTTCTAAGTACCTGGATTTC 181  
QY 186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205  
DB 182 ATCCATGTCATGACATATGACCTCCATGGCTCTCTGGAGGGCTACACTGGGGAGAAATAGT 241  
QY 206 ProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225  
DB 242 CCTCTTTCAAAATACCTCTACTGAGACTGGTGAATGCTTACCTACCTCAATGTGATTATGTC 301  
QY 226 MetAsnTyrTyrLysAsnGlnAlaProAlaGluLysLeuIleValGlyPheProGlu 245  
DB 302 ATGAATATTGGAAGAACATGAGAGCCAGCTGAGAGAGCTCATTTGGATTCCAGAG 361  
QY 246 TyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSer 265  
DB 362 TATGGACACACCTTCTCTGAGAAACCCCTCTGATAATGGAATTTGGTCCCTACCTCT 421  
QY 266 GlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTyrAlaTyrTyrGluIle 285  
DB 422 GGTGATGGCCCTGCTGGGCCCTATACACAGAGGCTGGGTTCTGGGCCCTACTATGAGATT 481  
QY 286 CysThrPheLeuArgSerGlyAlaThrGluValTyrAspAlaSerGlnGluValProTyr 305  
DB 482 TGCACCTTTCTGAGAGTGAGCCACTGAGGTCTGGGATGCTCCCAAGAGTGGCCCTAT 541  
QY 306 AlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325  
DB 542 GCCTATAGGCCCAACAGTGGCTTGGCTATGACATATCAAGAGCTTCAGTGTAAAGCT 601  
QY 326 GlnTyrLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTyrAlaIleAspLeuAspAsp 345  
DB 602 CAGTGGCTTAAGCAACAAATTTGGAGGTGCGATGCTGGGCCATTGAACCTTGATGAC 661  
QY 346 PheThrGlySerPheCysAspGlnGlyLys--PheProLeuThrSerThrLeuAsn--L 364  
DB 662 TTCCTGGCTCTTTCTGTGATCCAGGAACCAATTTCTCTTGACTTCTACTTTTGAACCAA 721  
QY 364 ysAlaLeuGlyIleSer-ThrGluGlyCysThrAlaProAspValProSerGluProVal 383  
DB 722 AAGCCCTTGGCATATCAACATCAAGGTTGCACAGCTCCCTGAGTGGCTTCCGAGCCAGTG 781  
QY 384 -ThrThr-ProProGlySerGlySerGlyGlySerSerGlyGlySerGlyGlys 403  
DB 782 AACTACTTCTCTCCAGGAAGTGGAGTGGGGGCTCGAAAAGCTCCGAGGGAAGCCCTTGA 841

QY 403 erGly-----PheCysAlaAspLysAlaAspGlyLeu-TyrProValAlaAspAsp 419  
DB 842 AGGCAGTCACTTCTGTCGCGCCAAAGCAGATGGCTACTAGCTGTGGCAGATGAC 901  
QY 420 ArgAsnAlaPheTyrGln 425  
DB 902 CC-AACTGCTTTTGGCAG 918

## RESULT 12

BG865815

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L14M10820 row: h column: 05

High quality sequence stop: 781.

Location/Qualifiers

1. 830

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone.lib="NCI CGAP SG2"

/lab host="DH10B (T1 phage-resistant)"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:

NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dr. Average insert size 1.3 Kb. Constructed by Life

Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 199 a 198 c 223 g 210 t

ORIGIN

Alignment Scores:

Pred. No.: 1 45e-137 Length: 830

Score: 1394.50 Matches: 269

Percent Similarity: 96.76% Conservativity: 0

Best Local Similarity: 96.76% Mismatches: 5

Query Match: 55.94% Indels: 6

DB: 12 Gaps: 0

US-10-004-219B-9 (1-452) x BG865815 (1-830)

QY 149 AlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyIleSer 168

DB 3 GCTATTGAGCAACAGCCAGACTGATGTTACTGCTGTAGCTGGTGGGATTTC 62

QY 169 AsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisVal 188

DB 63 AACATCCAGGCTGGCTATGAGATCCCTGAATTTCTAAGTACCTGATTCATCATGTC 122

QY 189 MetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyr 208

DB 123 ATGACATATGACCTCCATGGCTCTCTGGGAGGCTACACTGGGGAGAAATAGTCTCTTTAC 182

```
209 LysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyr 228
183 AAATACCTTACTGAGACTGGTAGCAATGCTACCTCAATGTGATTATGTCATGAATAT 242
229 TrpLysAsnGlyAlaProAlaGluLysLeuLeuValGlyPheProGluTyrGlyHis 248
243 TGAAGAACAATGGAGCCCGACGTGAGAGCTCATTTGTTGGATTCCAGATGATGGACAC 302
249 ThrPheLeuLeuArgAsnProSerAspAsnGlyLysGlyAlaProThrSerGlyAspGly 268
303 ACCTTCATCTCAGAAACCCCTCTGATATGAATGGTGGCCCTACCTCTCTGTGTGATGGC 362
269 ProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluLeuCysThrPhe 288
363 CTGTCTGGGCCCTATACACAGAGCGTGGTCTTGGCCCTACTATGAGATTTCACCTTT 422
289 LeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyrAlaTyrLys 308
423 CTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCCCAAGAGTGGCCCTATGCTATAG 482
309 AlaAsnGluTrpLeuGlyTyrAspAsnLeuLysSerPheSerValLysAlaGlnTrpLeu 328
483 GCCAAGAGTGGCTTGGCTATGACATATCAAGAGCTTCAGTGTAGGCTCAGTGGCTT 542
329 LysGlnAsnAsnPheGlyAlaMetIleTrpAlaIleAspLeuAspPheThrGly 348
543 AAGCAGAACAAATTTGGAGTGCCATGATCTGGCCATTGACCTTGATGACTTCACTGGC 602
349 SerPheCysAspGlnGly-LysPheProLeuThrSerThrLeuAsnLysAlaLeuGlyTl 368
603 TCTTTCTGTGATCAGGGAATAATTCCTCTGACTTCTACTTTGAACAAAGCCCTGGAT 662
368 eSerThrGluGlyCysThrAlaProAspValProSerGluProValThrProProgl 388
663 ATCCACTGAAGGTTCACAGCTCCTGACGTGCTCCGAGCAGTACTACTCTCCAGG 722
388 ySer--GlySerGlyGlySerSerGly-GlySerSerGlyGlySerGlyPheCysAl 407
723 AAAGTGGGAGTGGGGTGGAGTCCGAGGAGGAGCTCTGGAGGAGTGGATTCTGTGC 782
407 aAspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAla 422
783 CGACAAGCAGATGGCTC-TACCCTGT-TCAGATGACAGAATGGCT 826

RESULT 13
AV072529
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT
```

Fax: 81-45-503-9216  
Email: genome-res@gsr.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa  
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
Hayashizaki, Y.  
Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp/) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

## FEATURES

Location/Qualifiers  
1..1007  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="2200004B09"  
/sex="male"  
/tissue\_type="stomach"  
/dev\_stage="adult"  
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia  
RI; 1st strand cDNA was primed with a Not I - oligo(dT)  
primer [5'  
TGTACCAATCTGAGTGGAGCGCCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT73 vector.  
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M. Fatima  
Bonaldi."

BASE COUNT 248 a 246 c 243 g 270 t

## ORIGIN

Alignment Scores:  
Pred. No.: 4, 77e-136 Length: 1007  
Score: 1381.50 Matches: 269  
Percent Similarity: 95.44% Conservative: 3  
Best Local Similarity: 94.39% Mismatches: 12  
Query Match: 55.42% Indels: 6  
DB: 9 Gaps: 1

US-10-004-219B-9 (1-452) x AV072529 (1-1007)

QY 168 SerAsnIleGlnAlaGlyTyrGluLeuSerLysTyrLeuAspPheIleHis 187  
Db 60 TCCAACATCCAGCTCTATGAGAT---CCTTGACITTTCTAGTC-CTTGATTCTCATCAT 115  
QY 188 ValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeu 207  
Db 116 GTCATACCATATGACTTCTCCATGGCTCTCTGGGAGGCTACACTGGGAGAAATAGTCT-CTT 174



208 TyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsn 227  
Db : : : : :  
175 TATCAATACCTTACTGAGCTGGTAGCAATGCTTACTCAATGTGATATGTCATGAAC 234  
Qy  
228 TyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGly 247  
Db TATTGGAAGAACATAGGAGCCCGAGCTGAGAGCTCATTTGTTGA-TTCCAGAGATATGGA 293  
Qy  
248 HisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAsp 267  
Db CACACCTTCATCTCAGAAACCCCTCTAT-AATGAAATGT-GCCCTTACTCTGGTGTAT 351  
Qy  
268 GlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluIleCysThr 287  
Db GSCCTCTGCTGGSCCTATATACAGAGCGTGGTCTTGGCCCTACTATGAGATTTGACCC 411  
Qy  
288 PheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyrAlaTyr 307  
Db TTTCTGAGAAAGTGGAGCCACTGAGTCTGGGATGCTCCCAAGAAAGTGCCTATGCCTAT 471  
Qy  
308 LysAlaAsnGluTrpLeuGlyTyrPheAsnIleLysSerPheSerValLysAlaGlnTrp 327  
Db AAGGCCAACAGAGTGGCTTGGCTATGACAAATATCAAGAGCTTCAGTGTAAAGCTCAGTGG 531  
Qy  
328 LeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspPheThr 347  
Db CTTAAGCAGACAATTTTGGAGTGGCCATGATCTGGCCATTTGACCTTGATGACTTCCT 591  
Qy  
348 GlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAlaLeuGly 367  
Db GGCTCTTTCTGTGATCAGGAAATTTCTCTGACTTCTTACTTGAACAAAGCCCTTGCC 651  
Qy  
368 IleSerThrGluGlyCysThrAlaProAspValProSerGluProValThrProPro 387  
Db ATATCCACTGAAGTGGACAGCTCTGACGTCCTTCGAGCCCTTCGAGCCAGTACTCTCTCA 711  
Qy  
388 GlySerGlySerGlyGlyCysSerGlyGlySerSerGlyGlySerGlyPheCysAla 407  
Db GGAAGTGGAGTGGGGTGGAGCTCCGAGGAGCTCTGGAGGAGTGGATCTTGTCGC 771  
Qy  
408 AspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAlaPheTrpGlnCysIle 427  
Db GACAAAGCAGATGGCTCTTACCCTGGCAGATGACAGAAATGCTTTTGGCAGTGCATC 831  
Qy  
428 AsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPheAspThrSerCysAsn 447  
Db AATGGAATCACATACAGCAGCATTTGTCAAGCAGGCTTGTTTTGTATACCAAGCTGTAA 891  
Qy  
448 CysCysAsnTrpPro 452  
Db TGCTGCACTGGCCCA 906  
RESULT 14  
BG871434  
LOCUS  
DEFINITION 857 bp mRNA linear EST 29-MAY-2001  
602790636F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4921784 5',  
RNA sequence.  
ACCESSION BG871434  
VERSION BG871434.1 GI:14221974  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 857)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10840 row: h column: 09  
High quality sequence stop: 782.  
Location/Qualifiers  
1. 857  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4921784"  
/clone\_lib="NCI CGAP SG2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
Noti; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dr. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 222 a 203 c 214 g 218 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.42e-134 Length: 857  
Score: 1363.00 Matches: 273  
Percent Similarity: 95.50% Conservative: 3  
Best Local Similarity: 94.46% Mismatches: 8  
Query Match: 54.67% Indels: 8  
DB: 12 Gaps: 1  
US-10-004-219B-9 (1-452) x BG871434 (1-857)  
Qy 48 IleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSer 67  
Db 1 ATAGAAATGAATGATGTACTCTCTATTAAGCTTCAATGACTTGAATAACAGGACAGC 60  
Qy 68 LysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrAlaProPheThrThr 87  
Db 61 AAACCTGAAACCCCTCTCGCAATTTGGAGCTTGAACCTTTGGAACCTGCTTCTTCACTACC 120  
Qy 88 MetValSerThrSerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArg 107  
Db 121 ATGGTTTTCACCTTCTCAGAACCGCCAGACCTTCAATTAACCTCAGTCATCAAAATTTCTGGT 180  
Qy 108 GlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGlySerArgGlySerPro 127  
Db 181 CAGTATGGGTTTGTATGAGCTGGACCTGGAGCTGGGAATACCCAGGCTCAGTGGGAGCCCT 240  
Qy 128 ProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGln 147  
Db 241 CCTCAGGACAGCATCTCTTCACTGTCTGTGTGAAGAAATGCGTGAAGCTTTTGAGCAG 300  
Qy 148 GluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIle 167  
Db 301 GAGGCTATTGAGAGCAACAGGCCACAGCTGATGGTTACTGTCTGTGTAGCTGTGGGATTT 360  
Qy 168 SerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHis 187  
Db 361 TCCAAATCCAGGCTGGCTATGAGATCCCTGAACCTTCTTAAGTACCTGGAATTCATCCAT 420  
Qy 188 ValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeu 207  
Db 421 GTCATGACATATGACCTCCATGGCTCTCTGGAGGGCTACACCTGGGAGAAATAGTCTCTTT 480  
Qy 208 TyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsn 227  
Db 481 TACAAATACCTTACTTGAGCTGGTAGCAATGCTCACTCAATGTGGATTTATGTCATGAAC 540  
Qy 228 TyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGly 247  
Db 541 TATTGGAAGAACATAGGAGCCCGAGCTGAGAGCTCATTTGTTGATTTCCAGAGATATGA 600  
Qy 248 HisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAsp 267

Db 601 CACACCTTCATCTCGAGAACCTCT-GATAATGGAATTGGTCCCTACCTCTCGTGAT 659  
Qy 268 GlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTyr-AlaTyrTyrGlu-IleCysT 287  
Db 660 GGCCTTCGCTGGGCCCCATACAGACAGGCTGGGCTCTGGGGCTACTATGAGAAATTTGCA 719  
Qy 287 hPheLeuArgSer-GlyAlaThrGlu-ValTyrAspAlaSerGlnGluValProTyrAl 306  
Db 720 CTTTCTGAGAGTGGAGCCACTGAGGCTCTGGGATGA-CTCCAAGAGTGGCCTATGC 778  
Qy 306 aTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIle---LysSerPheSerValLysAl 325  
Db 779 CTATAAGGCCAAGAGTGGCTTCGATGATGACCC-ATATTCAAGAGCTTCAATGTTAAGGC 837  
Qy 325 aGlnTyrLeuLysGlnAsn 331  
Db 838 TCAGTGGGTTAAGCAGAAC 856

RESULT 15  
LOCUS BG866853  
DEFINITION 800 bp mRNA linear EST 29-MAY-2001  
ACCESSION 602786689F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4912438 5',  
VERSION BG866853 mRNA sequence.  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 800)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10816 row: b column: 23  
High quality sequence stop: 787.  
Location/Qualifiers  
1. 800  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4912438"  
/clone\_lib="NCI\_CGAP\_SG2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

## FEATURES

source  
194 a 190 c 205 g 211 t  
Location/Qualifiers  
1. 800  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4912438"  
/clone\_lib="NCI\_CGAP\_SG2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

## BASE COUNT

ORIGIN

## Alignment Scores:

Pred. No.: 3 54e-133 Length: 800  
Score: 1353.00 Matches: 257  
Percent Similarity: 97.36% Conservatives: 1  
Best Local Similarity: 96.98% Mismatches: 3  
Query Match: 54.27% Indels: 4  
DB: 12 Gaps: 0

US-10-004-219B-9 (1-452) x BG866853 (1-800)

Qy 104 LysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTyrProGlySer 123

Db 1 AAATTTCTGCGTCAATATGGGTTTGAATGGACCTGGACCTGGGAATACCCAGGCTCA 60  
Qy 124 ArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGlu 143  
Db 61 CTTGGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGCTCTGGTGAAGGAATGGTGA 120  
Qy 144 AlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaVal 163  
Db 121 GCTTTTGAAGCAGGAGCTATTGAGAGCAACAGGCCAGCATGATGGTTACTGCTCTGA 180  
Qy 164 AlaGlyGlyLysSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeu 183  
Db 181 GCTGTGGGATTTCAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCTTAAGTACCTG 240  
Qy 184 AspPheIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGlu 203  
Db 241 GATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTGGAGGGGTACACTGGGAG 300  
Qy 204 AsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAsp 223  
Db 301 AATAGTCTCTTTTCAAAATACCTACTGAGACTGGTAGCAATGCCCTACTCAATGGAT 360  
Qy 224 TyrValMetAsnTyrTyrLysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPhe 243  
Db 361 TATGTTCATGAACACTATTGGAAGAACAAATGGAGCCCGAGCTGAGAACTCATTTGGATT 420  
Qy 244 ProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla-Pr 263  
Db 421 CCAGAGTATGACACACCTTCTCTGAGAAACCCCTCTGATATGGAATGGTGGCCCCC 480  
Qy 263 oThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTyrAlaTyrTy 283  
Db 481 TACCTCTGGTGATGGCCCTGCTGGGCCCTATACACAGACAGAGCTGGGTTCTGGCCCTACTA 540  
Qy 283 rGluIleCysThrPheLeuArgSerGlyAlaThrGluValTyrAspAlaSerGlnGluVa 303  
Db 541 TGAGATTTGCACCTTTCTGAGAAGTGGAGCCCACTGAGGTCTGGGATGCTCCCAAGAAGT 600  
Qy 303 lProTyrAla-TyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerV 323  
Db 601 GCCCTATGCCCTATTAAGGCCAACGAGTGGCTTGGCTATGACATATCAAGAGCTTCAGTG 660  
Qy 323 aLysAlaGlnTyrLeuLysGlnAsnAsnPhe-GlyGlyAlaMetIleTyrAlaIleAsp 342  
Db 661 TTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGGCCATGATCTGGGCCATTGAA 720  
Qy 343 LeuAspAspPheThrGlySerPheCysAspGlnGly-LysPheProLeuThrSerThrIle 362  
Db 721 CTTGATGACTTCACTTGGTCTTTCTGTGATCAGGCCAAAAATTTCTCTGACTTCTACTTT 780  
Qy 362 uAsnLys 364  
Db 781 GAACAAG 787

Search completed: June 30, 2003, 01:07:36  
Job time : 1505.73 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:50:19 ; Search time 29.6228 Seconds  
(without alignments)  
2001.720 Million cell updates/sec

Title: US-10-004-219B-10

Perfect score: 2398

Sequence: 1 AKLVCFYTNWYQGEARF.....QQSCPTGLVFSNCKRCTWN 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	100.0	466	19 AAW40259	Human chitinase pr
2	2398	100.0	466	20 AAY42425	MO-218 clone of hu
3	2398	100.0	466	22 AAEO0432	Human chitinase pr
4	2398	100.0	466	23 ABB76291	Human chitinase.
5	2394	99.8	466	18 AAW08594	Human 50 kDa chiti
6	2392	99.7	466	19 AAW40260	Human chitinase pr
7	2392	99.7	466	20 AAY42426	MO-13B clone of hu
8	2392	99.7	466	22 AAEO0433	Human chitinase pr
9	2392	99.7	466	23 ABB76292	Human chitinase.
10	2391	99.7	466	18 AAW31498	Human chitotriosid

11	2372	98.9	464	18 AAW31497	Human chitotriosid
12	1982	82.7	373	19 AAW40261	Human chitinase pr
13	1982	82.7	373	20 AAY42427	Clone of the C-ter
14	1982	82.7	373	23 ABB76293	Human chitinase tr
15	1974	82.3	373	19 AAW40262	Human chitinase pr
16	1974	82.3	373	20 AAY42428	Chitinase amino ac
17	1974	82.3	373	23 ABB76294	Human chitinase an
18	1931	80.5	387	18 AAW08585	Human 39 kDa chiti
19	1333	55.2	455	22 AAG62541	Disease treatment
20	1333	55.2	476	22 AAG62543	Disease treatment
21	1053.5	43.9	421	19 AAW53122	Amino acid sequenc
22	1045.5	43.6	385	19 AAW53123	Amino acid sequenc
23	1045.5	43.6	385	19 AAW47033	Human cartilage gp
24	1045.5	43.6	416	19 AAW47034	Human cartilage gp
25	1045.5	43.6	421	19 AAW53121	Amino acid sequenc
26	1045.5	43.6	423	18 AAW35930	Human cartilage gl
27	1035.5	43.2	383	18 AAW26751	Bovine whey protei
28	1015.5	42.3	696	22 ABB58615	Drosophila melanog
29	1010.5	42.1	383	21 AAB03442	Gp38k protein sequ
30	981	40.9	368	23 ABB97580	Novel human protei
31	967.5	40.3	377	22 AAG62545	Disease treatment
32	967.5	40.3	398	22 AAG62544	Disease treatment
33	964.5	40.2	537	16 AAR73991	Bovine oviduct spe
34	942	39.3	718	16 AAR73992	Bovine oviduct spe
35	927	38.7	668	16 AAR73993	Hamster oviduct sp
36	919.5	38.3	4498	22 ABB58595	Drosophila melanog
37	848	35.4	460	22 ABB64366	Drosophila melanog
38	824	34.4	554	18 AAW01824	Manduca sexta larv
39	824	34.4	554	21 AAB07183	Manduca sexta gut
40	797	33.2	321	9 AAF81342	Polypeptide involv
41	786.5	32.8	595	22 ABB71737	Drosophila melanog
42	762	31.8	490	21 AAY52535	D. pteronyssius 98
43	762	31.8	490	23 AAU96339	Der HMW-map polype
44	762	31.8	509	21 AAY52533	D. pteronyssius 98
45	762	31.8	509	23 AAU96337	Der HMW-map polype

#### ALIGNMENTS

RESULT 1	
AAW40259	
ID AAW40259 standard; Protein; 466 AA.	
XX	
AC AAW40259;	
XX	
DT 15-JUN-1998 (first entry)	
XX	
DE Human chitinase protein from clone MO-218.	
XX	
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;	
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;	
KW rheumatoid arthritis; overexpression; extracellular matrix.	
XX	
OS Homo sapiens.	
XX	
FH Key	Location/Qualifiers
FT Peptide	1..21
FT Protein	/label= signal
FT FT	22..466
FT FT	/label= chitinase
FT FT	/note= "from clone MO-218"
XX	
PN WO9747752-A1.	
XX	
PD 18-DEC-1997.	
XX	
PF 16-JUN-1997; 97WO-US10460.	
XX	
DR 14-JUN-1996; 96US-0663618.	
XX	
PA (ICOS-) ICOS CORP.	
XX	

PI Gray PW;

XX WPI: 1998-052316/05.  
DR N-PSDB; AAV10435.

XX Nucleic acids encoding human chitinase - useful as antifungal  
PT agents, especially in combination with other antifungals

XX Claim 1; Page 41-42; 63pp; English.

XX This sequence represents a novel human chitinase isolated from clone  
CC MO-218. Chitinases are useful for treating or preventing fungal infection  
CC and as immunogens for generating antibodies which are used to purify,  
CC detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.  
CC The nucleic acid sequence of the chitinase is also useful as a probe to  
CC identify and isolate genomic DNA encoding chitinases or similar proteins,  
CC or cells expressing them or to generate transgenic ('knockout') rodents.  
CC It can also be used in hybridisation assays and to detect genetic  
CC alterations in the chitinase gene related to disease. Agents that  
CC this protein may be useful in treatment of Gaucher's disease and  
CC rheumatoid arthritis, where overexpression of the protein can damage  
CC the extracellular matrix. Chitinase also improves the activity of other  
CC antifungal agents and may allow a reduction in the dose of such agents,  
CC and thus of their side effects.

XX Sequence 466 AA;

Query Match 100.0%; Score 2398; DB 19; Length 466;  
Best Local Similarity 100.0%; Pred. No. 7.8e-221;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLVCFYTNWAQYRQGEARFLPKDPSLCTHLIYAFAGTNNHOLSTTWNDETLYQEFN 60

Db 22 AKLVCFYTNWAQYRQGEARFLPKDPSLCTHLIYAFAGTNNHOLSTTWNDETLYQEFN 81

Qy 61 GLKKNPKLTLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLRKYSFGDGLDWEY 120

Db 82 GLKKNPKLTLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLRKYSFGDGLDWEY 141

Qy 121 PGSQSPAVDKERTTLVQDLANAFQEAQTSQGERLLLSAIVPAGQYVDAGYEVNDKIA 180

Db 142 PGSQSPAVDKERTTLVQDLANAFQEAQTSQGERLLLSAIVPAGQYVDAGYEVNDKIA 201

Qy 181 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKROESGAASLNVDAAVOQLQKGPASKLI 240

Db 202 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKROESGAASLNVDAAVOQLQKGPASKLI 261

Qy 241 LGMTYGRSFTLASSSDTRVGAPATGSGTPGPFKEGGMLAYEYVCSWKGATKQRIQDQK 300

Db 262 LGMTYGRSFTLASSSDTRVGAPATGSGTPGPFKEGGMLAYEYVCSWKGATKQRIQDQK 321

Qy 301 VPYIFRDQNWGFDDVSFKTKVSYLKQKGLGGAMWALDLDFFAGSCNQRYPLOTL 360

Db 322 VPYIFRDQNWGFDDVSFKTKVSYLKQKGLGGAMWALDLDFFAGSCNQRYPLOTL 381

Qy 361 RQELSLPVLPGSTPELEVPKQPSSEPHGSPSGQDTFCQKADGLYPNPRSSFYSCA 420

Db 382 RQELSLPVLPGSTPELEVPKQPSSEPHGSPSGQDTFCQKADGLYPNPRSSFYSCA 441

Qy 421 AGRLFQOSCPGLVFSNSCKCTWN 445

Db 442 AGRLFQOSCPGLVFSNSCKCTWN 466

RESULT 2

AAV42425

ID AAV42425 standard; Protein; 466 AA.

XX AA42425;

XX 10-DEC-1999 (first entry)

XX MO-218 clone of human Chitinase, amino acid sequence.

XX

KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;  
KW organ transplant; parasite; chitin-binding; allele; vector;  
KW truncated protein; chitin binding region.

XX Homo sapiens.

XX Key Location/Qualifiers

PH Peptide 1..21

FT Protein /label= Signal\_peptide

FT Protein 22..445

FT Protein /label= Mature\_protein

FT Peptide /note= "Human Chitinase"

FT Peptide 391..445

FT Peptide /label= Chitin binding

FT Peptide /note= "Chitin binding region"

XX WO9946390-A1.

XX 16-SEP-1999.

XX 12-MAR-1999; 99WO-US05343.

XX 12-MAR-1998; 98US-0039198.

XX (ICOS-) ICOS CORP.

XX Gray PW, Tjoelker LW;

XX WPI; 1999-551417/46.

XX N-PSDB; AAZ21847.

XX Novel chitin-binding fragments of human chitinase used to treat fungal  
PT infections in animals

XX Example 1; Page 58-59; 83pp; English.

XX This is the amino acid sequence of an allelic form of the human  
CC chitinase enzyme, which is capable of degrading Chitin (a linear  
CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).  
CC The sequence also includes a Chitin binding peptide region, which lacks  
CC any chitinase activity, but which can be used to target anti-fungal  
CC agents to chitin containing fungal cells.

CC Chitinase fragments can be used to screen for proteins or other  
CC molecules that specifically bind to the chitin-binding domain of human  
CC chitinase or that modulate its activity. These compounds are useful for  
CC immunization, as well as for purifying chitinase, as well as for  
CC detection and quantification of chitinase. Polynucleotide fragments of  
CC the invention are useful as a source of probes and primers, and to  
CC express the proteins recombinantly. The chitinase fragments, when  
CC conjugated to antifungal compounds, are used to treat animals.

CC especially humans, infected with chitin-containing parasites such as  
CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
CC sporotrichosis, and dermatophytosis.

CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
CC protein for treating infections, especially fungal infections, is  
CC problematic. In view of the increasing incidents of life-threatening  
CC fungal infection in e.g. immunocompromised individuals, there exists a  
CC need for identifying new compounds for treating fungal infection. The  
CC chitin-binding fragments of the present invention provide this need.

XX Sequence 466 AA;

Query Match 100.0%; Score 2398; DB 20; Length 466;

Best Local Similarity 100.0%; Pred. No. 7.8e-221;

Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLVCFYTNWAQYRQGEARFLPKDPSLCTHLIYAFAGTNNHOLSTTWNDETLYQEFN 60

Db 22 AKLVCFYTNWAQYRQGEARFLPKDPSLCTHLIYAFAGTNNHOLSTTWNDETLYQEFN 81

QY 61 GLKONPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDLDWEY 120  
DB |||||  
82 GLKONPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDLDWEY 141  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQQAQTSKGRLLLSAAVPAGQTVVDAGYVDKIA 180  
DB |||||  
142 PGSQSPAVDKERFTTLVQDLANAFQQAQTSKGRLLLSAAVPAGQTVVDAGYVDKIA 201  
QY 181 QNLDFVNLMAVDHFGSWKVTGHNPLKRYKQESGAASLNVDAAVQWLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDHFGSWKVTGHNPLKRYKQESGAASLNVDAAVQWLQKGTTPASKLI 261  
QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKGAQKRIQDOK 300  
DB 262 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKGAQKRIQDOK 321  
QY 301 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGGAMVWALDLDFFAGFSNCGRYPLIOTL 360  
DB 322 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGGAMVWALDLDFFAGFSNCGRYPLIOTL 381  
QY 361 ROELSLPYLPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRERSFYSCA 420  
DB 382 ROELSLPYLPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRERSFYSCA 441

## RESULT 3

AAE00432  
ID AAE00432 standard; Protein; 466 AA.  
AC AAE00432;  
XX  
DT 19-JUN-2001 (first entry)  
XX Human chitinase protein from clone pMO-218.  
DE  
XX Human; antifungal; chitinase; immunoglobulin; Ig; therapy;  
KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;  
KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;  
KW clone pMO-218.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /label= Signal\_peptide  
FT Protein 22..466  
FT /label= Human\_mature\_chitinase\_protein  
FT Domain 418..466  
FT /label= Chitin\_binding\_domain  
FT Region 1..373  
FT /note= "Region with triacetylchitotriose hydrolysing activity"  
XX  
PN WO200123430-A2.  
XX  
XX 05-APR-2001.  
PD  
XX 28-SEP-2000; 2000WO-US26960.  
PF  
XX 30-SEP-1999; 99US-0409918.  
PR  
XX (ICOS-) ICOS CORP.  
PA Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;  
XX  
XX WPI; 2001-266141/27.  
DR N-PSDB; AAD03759.  
XX

PT Novel chitinase immunoglobulin fusion product, useful for treating  
PT fungal infections and reducing the amount of a non-chitinase antifungal  
PT agent needed for the treatment -  
XX  
XX Claim 1; Page 32-33; 39pp; English.  
PS  
XX The present invention relates to a chitinase immunoglobulin (Ig) fusion  
CC product, comprising a human chitinase fused to at least a portion of an  
CC immunoglobulin chain. The fusion product is useful for treating fungal  
CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,  
CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.  
CC The fusion protein is useful for reducing the amount of non-chitinase  
CC antifungal agent needed to exert an antifungal activity. The fusion  
CC protein is also useful for preparing  
CC a medicament for the prophylactic or therapeutic treatment of fungal  
CC infections. Chitinase immunoglobulin fusion product has unexpectedly  
CC improved serum half-life and formulation fusion properties.  
CC The present sequence is human chitinase protein from clone pMO-218.  
CC Chitinase enzyme degrades chitin which is a homopolymer of  
CC beta-(1,4)-linked N-acetylglucosamine residues.  
XX  
SQ Sequence 466 AA;

Query Match 100.0%; Score 2398; DB 22; Length 466;  
Best Local Similarity 100.0%; Pred. No. 7.8e-221;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLDPISLCYTHLIYAFAGMTHNQLSTTEWNETLYQEFN 60  
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLDPISLCYTHLIYAFAGMTHNQLSTTEWNETLYQEFN 81  
QY 61 GLKONPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDLDWEY 120  
DB 82 GLKONPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDLDWEY 141  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQQAQTSKGRLLLSAAVPAGQTVVDAGYVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVQDLANAFQQAQTSKGRLLLSAAVPAGQTVVDAGYVDKIA 201  
QY 181 QNLDFVNLMAVDHFGSWKVTGHNPLKRYKQESGAASLNVDAAVQWLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDHFGSWKVTGHNPLKRYKQESGAASLNVDAAVQWLQKGTTPASKLI 261  
QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKGAQKRIQDOK 300  
DB 262 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKGAQKRIQDOK 321  
QY 301 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGGAMVWALDLDFFAGFSNCGRYPLIOTL 360  
DB 322 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGGAMVWALDLDFFAGFSNCGRYPLIOTL 381  
QY 361 ROELSLPYLPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRERSFYSCA 420  
DB 382 ROELSLPYLPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRERSFYSCA 441  
QY 421 AGRLFQSQSCTGLVFSNSCKCCTWN 445  
DB 442 AGRLFQSQSCTGLVFSNSCKCCTWN 466

## RESULT 4

ABB76291  
ID ABB76291 standard; Protein; 466 AA.  
XX  
AC ABB76291;  
XX  
DT 12-AUG-2002 (first entry)  
XX Human chitinase.  
XX Chitinase; enzyme; human; fungicide; antifungal; infection;  
KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;

KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
XX Pneumocystis.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT Protein /label= Signal\_peptide  
FT Protein 22..466  
FT Protein /label= Mature\_protein  
XX  
PN US6372212-B1.  
XX  
XX 16-APR-2002.  
PD 16-JUN-1997; 97US-0877599.  
PF 14-JUN-1996; 96US-0663618.  
PR (ICOS-) ICOS CORP.  
PA  
XX Gray PW;  
PI  
XX  
DR WPI; 2002-442449/47.  
DR N-PSDB; ABL57380.  
XX  
XX Co-administering chitinase to improve the effectiveness of fungicidal  
PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -  
XX  
XX Claim 1; Column 27-30; 26pp; English.  
XX  
XX The present sequence is the protein sequence of human chitinase  
CC as predicted from isolated cDNA clone MO-218 (see ABL57380). It  
CC differs in only 1 amino acid residue from the chitinase sequence  
CC (see ABB76292) deduced from a second cDNA clone, having glycine  
CC at position 81 of the mature protein. Northern blots showed  
CC highest chitinase gene expression in lung and ovary tissues.  
CC Expression in lung is consistent with a protective role against  
CC pathogenic organisms that contain chitin. The invention provides  
CC human chitinase polynucleotides and polypeptides, and materials and  
CC methods for the recombinant production of human chitinase products,  
CC which are expected to be useful as products for treating fungal  
CC infections or for the development of such products. Human  
CC chitinase has a synergistic effect on the actions of other  
CC fungicides. It can be administered to improve the antifungal  
CC activity of a non-chitinase antifungal agent, especially  
CC amphotericin B or itraconazole, in the treatment of a fungal  
CC infection such as candidiasis, aspergillosis, coccidioidomycosis,  
CC blastomycosis, paracoccidioidomycosis, histoplasmosis,  
CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,  
CC dermatophytoses and Pneumocystis infections (all claimed). In  
CC particular, the fungal infection involves Candida, Aspergillus  
CC and/or Cryptococcus spp., whose growth is not effectively  
CC inhibited by contact with human chitinase alone.  
XX  
XX Sequence 466 AA;  
SQ

Query Match 100.0%; Score 2398; DB 23; Length 466;  
Best Local Similarity 100.0%; Pred. No. 7.8e-221;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKLVCFYTNWAOYRQGEARLPKLDPSLCTHLIYAPAGTNNHQLSTWENDETLTYQEFN 60  
Db |||||  
22 AKLVCFYTNWAOYRQGEARLPKLDPSLCTHLIYAPAGTNNHQLSTWENDETLTYQEFN 81  
QY 61 GLKKMNPKLTLAIGWNEFTQKFTDMVATANNROTFFNSAIRFLRKYSFDGLDLWWEY 120  
Db |||||  
82 GLKKMNPKLTLAIGWNEFTQKFTDMVATANNROTFFNSAIRFLRKYSFDGLDLWWEY 141  
QY 121 PGSQGSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAQGYVDAGYEVDKIA 180  
|||

Db 142 PGSQGSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAQGYVDAGYEVDKIA 201  
QY 181 QNLDFVNLWAYDFHGSWEKVTGHNSPLYKROESGAAASLNVDAAVQOQWLQKGTTPASKLI 240  
Db |||||  
202 QNLDFVNLWAYDFHGSWEKVTGHNSPLYKROESGAAASLNVDAAVQOQWLQKGTTPASKLI 261  
QY 241 LGMPTYGRSFTLLASSSDTRVGAPATGSGTGPFTKEGGMLAYEYVCSWKGATKQRIODQK 300  
Db |||||  
262 LGMPTYGRSFTLLASSSDTRVGAPATGSGTGPFTKEGGMLAYEYVCSWKGATKQRIODQK 321  
QY 301 VPYIFRDQWVGFDVSEFKTKVYLKQKGLGAMWALDLDLDDFAGFSCNQGRIYPLIQT 360  
Db |||||  
322 VPYIFRDQWVGFDVSEFKTKVYLKQKGLGAMWALDLDLDDFAGFSCNQGRIYPLIQT 381  
QY 361 ROELSPLPSTGTPPELVKPGQSPPEHGPSPGQDTFCQKADGLYPNPRSSSFYSKA 420  
Db |||||  
382 ROELSPLPSTGTPPELVKPGQSPPEHGPSPGQDTFCQKADGLYPNPRSSSFYSKA 441  
QY 421 AGRLFQOQSCPTGLVFSNSCKCCTWN 445  
Db |||||  
442 AGRLFQOQSCPTGLVFSNSCKCCTWN 466  
XX  
XX RESULT 5  
XX AAW08584  
ID AAW08584 standard; Protein; 466 AA.  
XX  
XX AC AAW08584;  
XX DT 24-MAR-1997 (first entry)  
XX DE Human 50 kDa chitinase.  
XX KW Chitinase; chitotriosidase; chitin; infectious disease;  
KW gene therapy; vaccine; diagnosis; lysosomal lipidosis;  
KW Gaucher disease; leishmaniasis; sarcoidosis; multiple sclerosis;  
XX X-linked adrenoleukodystrophy; drug delivery; cosmetics; food.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..21  
FT Domain /label= Sig\_peptide  
FT Domain 178..198  
FT FT /label= Catalytic\_domain  
XX WO9640940-A2.  
PN 19-DEC-1996.  
XX 06-JUN-1996; 96WO-NL00225.  
XX 07-JUN-1995; 95US-0486839.  
PR (UNAM ) UNIV AMSTERDAM.  
XX Aerts JWFG;  
XX WPI; 1997-118698/11.  
DR N-PSDB; AAT50833.  
XX  
XX New human chitinase - used to treat or prevent infection by  
PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,  
PT foods, implants etc.  
XX  
PS Claim 1; Page 40-42; 58pp; English.  
XX  
CC Human chitinases of 50 kDa (AAW08584) and 39 kDa (AAW08585) are stable  
CC to many proteases, active at pH 3-8 and up to 50 deg, and stable in  
CC the circulation. They are the result of alternative splicing of  
CC RNA, and can be produced on a large-scale in transformed host cells  
CC using cDNA clones (see also AAT50833-34) obtd. from a human macrophage  
CC library. The chitinases, optionally expressed from a gene therapy

CC vector, are used to treat or prevent infection by chitin-contg.  
CC pathogens (e.g. fungi, protozoa, helminths) and more generally to  
CC degrade chitin. They are also useful in controlled-release drug  
CC delivery, in cosmetics, foods and dental products, for antibody  
CC prodn. and for diagnosis of diseases associated with elevated  
CC chitinase levels.

XX SQ Sequence 466 AA;

Query Match 99.8%; Score 2394; DB 18; Length 466;  
Best Local Similarity 100.0%; Pred. No. 1.9e-220;  
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KLVVYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYQEFNG 61  
DB 23 KLVVYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYQEFNG 82  
QY 62 LKQWNPCLKTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEYP 121  
DB 83 LKQWNPCLKTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEYP 142  
QY 122 GSQGSFPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYVDKIAQ 181  
DB 143 GSQGSFPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYVDKIAQ 202  
QY 182 NLDFVNLMAVDPHGSEKVTGHNPSPLYKROESGAAASLNVDAAVQWMLQKGTTPASKLIL 241  
DB 203 NLDFVNLMAVDPHGSEKVTGHNPSPLYKROESGAAASLNVDAAVQWMLQKGTTPASKLIL 262  
QY 242 GMPTVGRSFTLASSSDTRVGAPATSGTGPPTKEGGMLAYEVCVSKGATKQRTQDQKV 301  
DB 263 GMPTVGRSFTLASSSDTRVGAPATSGTGPPTKEGGMLAYEVCVSKGATKQRTQDQKV 322  
QY 302 PYIFRDQWGFDDVESFKTKVSYLKQGLGAMVWALDLDFFAGSCNQGRIPLIQTLR 361  
DB 323 PYIFRDQWGFDDVESFKTKVSYLKQGLGAMVWALDLDFFAGSCNQGRIPLIQTLR 382  
QY 362 QELSPLYPSTGTELEVPKPGPSEPEHGFSPGQDTFCQKADGLYPNPRRSPFYSCAA 421  
DB 383 QELSPLYPSTGTELEVPKPGPSEPEHGFSPGQDTFCQKADGLYPNPRRSPFYSCAA 442  
QY 422 GRLFOQSCPTGLVFSNSCKCCTWN 445  
DB 443 GRLFOQSCPTGLVFSNSCKCCTWN 466

RESULT 6  
AAW40260  
ID AAW40260 standard; Protein; 466 AA.

XX AC AAW40260;  
XX DT 15-JUN-1998 (first entry)  
XX DE Human chitinase protein from clone MO-13B.

XX KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;  
XX KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;  
XX KW rheumatoid arthritis; overexpression; extracellular matrix.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT Peptide 1..21  
XX FT 22..466 /label= signal  
XX FT Protein /label= chitinase  
XX FT /note= "from clone MO-13B"

XX PN W09747752-A1.

XX PD 18-DEC-1997.

PF 16-JUN-1997; 97WO-US10460.  
XX 14-JUN-1996; 96US-0663618.  
XX (ICOS-) ICOS CORP.  
XX PI Gray PW;  
XX WPI; 1998-052316/05.  
XX N-PSDB; AAV10436.  
XX Nucleic acids encoding human chitinase - useful as antifungal  
XX agents, especially in combination with other antifungals  
XX Claim 7; Page 44-45; 63pp; English.

CC This sequence represents a novel human chitinase isolated from clone  
CC MO-13B. Chitinases are useful for treating or preventing fungal infection  
CC and as immunogens for generating antibodies which are used to purify,  
CC detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.  
CC The nucleic acid sequence of the chitinase is also useful as a probe to  
CC identify and isolate genomic DNA encoding chitinases or similar proteins,  
CC or cells expressing them or to generate transgenic ('knockout') rodents.  
CC It can also be used in hybridisation assays and to detect genetic  
CC alterations in the chitinase gene related to disease. Agents that inhibit  
CC this protein may be useful in treatment of Gaucher's disease and  
CC rheumatoid arthritis, where overexpression of the protein can damage  
CC the extracellular matrix. Chitinase also improves the activity of other  
CC antifungal agents and may allow a reduction in the dose of such agents,  
CC and thus of their side effects.

XX SQ Sequence 466 AA;

Query Match 99.7%; Score 2392; DB 19; Length 466;  
Best Local Similarity 99.8%; Pred. No. 2.9e-220;  
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYQEFN 60  
DB 22 AKLVYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYQEFN 81  
QY 61 GLKQWNPCLKTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEY 120  
DB 82 GLKQWNPCLKTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEY 141  
QY 121 PGSQGSFPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYVDKIA 180  
DB 142 PGSQGSFPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYVDKIA 201  
QY 181 QNLDFVNLMAVDPHGSEKVTGHNPSPLYKROESGAAASLNVDAAVQWMLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDPHGSEKVTGHNPSPLYKROESGAAASLNVDAAVQWMLQKGTTPASKLI 261  
QY 241 LGMPYGRSFTLASSSDTRVGAPATSGTGPPTKEGGMLAYEVCVSKGATKQRTQDQK 300  
DB 262 LGMPYGRSFTLASSSDTRVGAPATSGTGPPTKEGGMLAYEVCVSKGATKQRTQDQK 321  
QY 301 VPIYIFRDQWGFDDVESFKTKVSYLKQGLGAMVWALDLDFFAGSCNQGRIPLIQT 360  
DB 322 VPIYIFRDQWGFDDVESFKTKVSYLKQGLGAMVWALDLDFFAGSCNQGRIPLIQT 381  
QY 361 ROELSPLYPSTGTELEVPKPGPSEPEHGFSPGQDTFCQKADGLYPNPRRSPFYSCA 420  
DB 382 ROELSPLYPSTGTELEVPKPGPSEPEHGFSPGQDTFCQKADGLYPNPRRSPFYSCA 441  
QY 421 AGRLFQOQSCPTGLVFSNSCKCCTWN 445  
DB 442 AGRLFQOQSCPTGLVFSNSCKCCTWN 466

RESULT 7  
AA42426  
ID AAY42426 standard; Protein; 466 AA.

XX AC AAY42426;  
 XX DT 10-DEC-1999 (first entry)  
 XX DE MO-13B clone of human Chitinase, amino acid sequence.  
 XX KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;  
 KW organ transplant; parasite; chitin-binding; allele; vector;  
 KW truncated protein; chitin binding region.  
 XX OS Homo sapiens.  
 XX FH Key  
 FT Peptide 1..21  
 FT Protein /label= Signal\_peptide  
 FT /label= Mature\_protein  
 FT WO9946390-A1.  
 XX PN 16-SEP-1999.  
 XX PD 12-MAR-1999; 99WO-US05343.  
 XX PR 12-MAR-1998; 98US-0039198.  
 XX PA (ICOS-) ICOS CORP.  
 XX PI Gray PW, Tjoelker LW;  
 XX WPI; 1999-551417/46.  
 XX DR N-PSDB; AAZ21848.  
 XX PT Novel chitin-binding fragments of human chitinase used to treat fungal  
 XX infections in animals -  
 XX PS Example 1; Page 62-64; 83pp; English.  
 XX CC This is the amino acid sequence of an allelic form of the human  
 CC chitinase enzyme, which is capable of degrading Chitin (a linear  
 CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).  
 CC The sequence also includes a Chitin binding peptide region, which lacks  
 CC any chitinase activity, but which can be used to target anti-fungal  
 CC agents to chitin containing fungal cells.  
 CC Chitinase fragments can be used to screen for proteins or other  
 CC molecules that specifically bind to the chitin-binding domain of human  
 CC chitinase or that modulate its activity. These compounds are useful for  
 CC immunization, as well as for purifying chitinase, as well as for  
 CC detection and quantification of chitinase. Polynucleotide fragments of  
 CC the invention are useful as a source of probes and primers, and to  
 CC express the proteins recombinantly. The chitinase fragments, when  
 CC conjugated to antifungal compounds, are used to treat animals,  
 CC especially humans, infected with chitin-containing parasites such as  
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
 CC sporotrichosis, and dermatophytoses.  
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
 CC protein for treating infections, especially fungal infections, is  
 CC problematic. In view of the increasing incidents of life-threatening  
 CC fungal infection in e.g. immunocompromised individuals, there exists a  
 CC need for identifying new compounds for treating fungal infection. The  
 CC chitin-binding fragments of the present invention provide this need.  
 XX SQ Sequence 466 AA;  
 Query Match 99.7%; Score 2392; DB 20; Length 466;  
 Best Local Similarity 99.8%; Pred. No. 2.9e-220;  
 Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 1 AKLVCFYTNWAOYRQGEARFLPKDLPDLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 60  
 |||

DB 22 AKLVCFYTNWAOYRQGEARFLPKDLPDLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 81  
 QY 61 GLKKNPKLKTLLAIGGWNFGTKETDMVATANNRQTFVNSAIRFLRKYSFDGLDLDEY 120  
 DB 82 GLKKNPKLKTLLAIGGWNFGTKETDMVATANNRQTFVNSAIRFLRKYSFDGLDLDEY 141  
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQQAQTSKGERLLLSAAVPAQOTVVDAGYEVDKIA 180  
 DB 142 PGSQSPAVDKERFTTLVQDLANAFQQAQTSKGERLLLSAAVPAQOTVVDAGYEVDKIA 201  
 QY 181 QNLDFNLMAVDYFHGSWEKVTGHNPSPLYKROESGAASLNVDAAVQOVLQKGTASKLI 240  
 DB 202 QNLDFNLMAVDYFHGSWEKVTGHNPSPLYKROESGAASLNVDAAVQOVLQKGTASKLI 261  
 QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKATQRIQDOK 300  
 DB 262 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKATQRIQDOK 321  
 QY 301 VPIYIFRDNQWVGFDVESFKTKVSYLKQKLGAMVWALDLDLDDFAGFSCNQGRYPLIQTL 360  
 DB 322 VPIYIFRDNQWVGFDVESFKTKVSYLKQKLGAMVWALDLDLDDFAGFSCNQGRYPLIQTL 381  
 QY 361 ROELSLPYLPSGTPLEVPKPGQSPGPHGSPGODTFCQKADGLYNPNRERSFYSCA 420  
 DB 382 ROELSLPYLPSGTPLEVPKPGQSPGPHGSPGODTFCQKADGLYNPNRERSFYSCA 441  
 QY 421 AGRLFQOQSCPTGLVFSNCKCCTWN 445  
 DB 442 AGRLFQOQSCPTGLVFSNCKCCTWN 466  
 RESULT 8  
 ID AAE00433 standard; Protein; 466 AA.  
 XX AC AAE00433;  
 XX DT 19-JUN-2001 (first entry)  
 XX DE Human chitinase protein from clone pMO-13B.  
 XX KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;  
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;  
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;  
 KW clone pMO-13B.  
 XX OS Homo sapiens.  
 XX FH Key  
 FT Peptide 1..21  
 FT Protein /label= Signal\_peptide  
 FT /label= Human\_mature\_chitinase\_protein  
 FT Domain 418..466  
 FT /label= Chitin\_binding\_domain  
 FT Region 1..373  
 FT /note= "Region with triacetylechitotriose  
 FT hydrolysing activity"  
 XX PN WO2001213430-A2.  
 XX PD 05-APR-2001.  
 XX PF 28-SEP-2000; 2000WO-US26960.  
 XX PR 30-SEP-1999; 99US-0409918.  
 XX PA (ICOS-) ICOS CORP.  
 XX PI Allison DS, Dietech GN, Gray PW, Shaw KD, Steiner BH;  
 XX WPI; 2001-266141/27.



DR N-PSDB; AAD03760.  
XX Novel chitinase immunoglobulin fusion product, useful for treating  
PT fungal infections and reducing the amount of a non-chitinase antifungal  
PT agent needed for the treatment -  
XX  
XX Claim 1; Page 36-38; 39pp; English.  
XX  
XX The present invention relates to a chitinase immunoglobulin (Ig) fusion  
CC product, comprising a human chitinase fused to at least a portion of an  
CC immunoglobulin chain. The fusion product is useful for treating fungal  
CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,  
CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.  
CC The fusion protein is useful for reducing the amount of non-chitinase  
CC antifungal agent needed to exert an antifungal activity. The fusion  
CC protein is also useful for preparing  
CC a medicament for the prophylactic or therapeutic treatment of fungal  
CC infections. Chitinase immunoglobulin fusion product has unexpectedly  
CC improved serum half-life and formulation properties.  
CC The present sequence is human chitinase protein from clone pMO-13B.  
CC Chitinase enzyme degrades chitin which is a homopolymer of  
CC beta-(1,4)-linked N-acetylglucosamine residues.  
XX  
SQ Sequence 466 AA;

Query Match 99.7%; Score 2392; DB 22; Length 466;  
Best Local Similarity 99.8%; Pred. No. 2.9e-220;  
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQLSTTEWDETLYQSFN 60  
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQLSTTEWDETLYQSFN 81  
QY 61 GLKXNPKLKTLLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEY 120  
DB 82 GLKXNPKLKTLLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEY 141  
QY 121 PGSQSPAVDKERFTLLVQDLANAFQQAQTSKGERLLLSAAVPAQTVVDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTLLVQDLANAFQQAQTSKGERLLLSAAVPAQTVVDAGYEVDKIA 201  
QY 181 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKQESGAASLNVDAAVQWLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKQESGAASLNVDAAVQWLQKGTTPASKLI 261  
QY 241 LGMPYGRSFTLASSDTRVGAPATGSGTPGFTKEGGLAYEYVCSWKGATKQRIQDK 300  
DB 262 LGMPYGRSFTLASSDTRVGAPATGSGTPGFTKEGGLAYEYVCSWKGATKQRIQDK 321  
QY 301 VPIYIFRDQWVGFDVGEFKTKVSVYLKQGLGAMVWALDLDFFAGFCNQGRIYPLIOTL 360  
DB 322 VPIYIFRDQWVGFDVGEFKTKVSVYLKQGLGAMVWALDLDFFAGFCNQGRIYPLIOTL 381  
QY 361 ROELSLPLPSGTPELVKPKQPSPEHPGSPGQDTFCQKADGLYNPNRERSFYSCA 420  
DB 382 ROELSLPLPSGTPELVKPKQPSPEHPGSPGQDTFCQKADGLYNPNRERSFYSCA 441  
QY 421 AGRLFQSCPTGLVFSNCKCCTWN 445  
DB 442 AGRLFQSCPTGLVFSNCKCCTWN 466

RESULT 9  
ABB76292  
ID ABB76292 standard; Protein; 466 AA.

XX  
AC ABB76292;

XX  
DT 12-AUG-2002 (first entry)

XX  
DE Human chitinase.

XX

KW Chitinase; enzyme; human; fungicide; antifungal; infection;  
KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;  
KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
KW Pneumocystis.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Peptide 1..21  
XX Protein /label= Signal\_peptide  
XX /label= Mature\_protein  
XX US6372212-B1.  
XX 16-APR-2002.  
XX 16-JUN-1997; 97US-0877599.  
XX 14-JUN-1996; 96US-0663618.  
XX (ICOS-) ICOS CORP.  
XX Gray PW;  
XX WPI; 2002-442449/47.  
XX N-PSDB; ABL57381.

Co-administering chitinase to improve the effectiveness of fungicidal  
drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
infections e.g. candidiasis, coccidioidomycosis and blastomycosis -

Example 1; Column 33-36; 26pp; English.

The present sequence is the protein sequence of human chitinase  
as predicted from isolated cDNA clone MO-13B (see ABL57378). It  
differs in only 1 amino acid residue from the chitinase sequence  
(see ABB76291) deduced from a second cDNA clone, having serine  
at position 81 of the mature protein. Northern blots showed  
highest chitinase gene expression in lung and ovary tissues.  
Expression in lung is consistent with a protective role against  
pathogenic organisms that contain chitin. The invention provides  
human chitinase polynucleotides and polypeptides, and materials and  
methods for the recombinant production of human chitinase products,  
which are expected to be useful as products for treating fungal  
infections or for the development of such products. Human  
chitinase has a synergistic effect on the actions of other  
fungicides. It can be administered to improve the antifungal  
activity of a non-chitinase antifungal agent, especially  
amphotericin B or itraconazole, in the treatment of a fungal  
infection such as candidiasis, aspergillosis, coccidioidomycosis,  
blastomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
chromoblastomycosis, sporotrichosis, dermatophytosis, and  
dermatophytoses and Pneumocystis infections (all claimed). In  
particular, the fungal infection involves Candida, Aspergillus  
and/or Cryptococcus spp., whose growth is not effectively  
inhibited by contact with human chitinase alone.

Query Match 99.7%; Score 2392; DB 23; Length 466;  
Best Local Similarity 99.8%; Pred. No. 2.9e-220;  
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQLSTTEWDETLYQSFN 60  
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQLSTTEWDETLYQSFN 81  
QY 61 GLKXNPKLKTLLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEY 120  
DB 82 GLKXNPKLKTLLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEY 141

QY 121 PGSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIA 180  
DB 142 PGSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIA 201  
QY 181 QNLDFVNLWAYDFHGSWEKVTGHSNPLYKRBESGAAASLNVDAAVQWLOKGTTPASKLI 240  
DB 202 QNLDFVNLWAYDFHGSWEKVTGHSNPLYKRBESGAAASLNVDAAVQWLOKGTTPASKLI 261  
QY 241 LGMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYVEVCSWGKATKORIQDOK 300  
DB 262 LGMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYVEVCSWGKATKORIQDOK 321  
QY 301 VPIYFRDNQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRYPLIQTL 360  
DB 322 VPIYFRDNQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRYPLIQTL 381  
QY 361 ROELSPLPYLPSGTPPELVKPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 420  
DB 382 ROELSPLPYLPSGTPPELVKPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 441  
QY 421 AGRLFQSCPTGLVFSNSCKCCTWN 445  
DB 442 AGRLFQSCPTGLVFSNSCKCCTWN 466

RESULT 10

AAW31498  
ID AAW31498 standard; Protein; 466 AA.

AC AAW31498;

XX 27-APR-1998 (first entry)

DE Human chitotriosidase variant.

XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;  
KW rheumatoid arthritis; atherosclerosis; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 102

FT /note= "encoded by RGC"

XX WO9736917-A1.

XX 09-OCT-1997.

XX 21-MAR-1997; 97WO-US05072.

XX 29-MAR-1996; 96US-0014295.

XX (HUMA-) HUMAN GENOME SCI INC.  
PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;

XX WPI; 1997-503041/46.

DR N-PSDB; AAT89181.

XX New isolated human chitotriosidase gene - used to develop products  
PT for the diagnosis and treatment of tissue remodeling disorders, e.g.  
PT rheumatoid arthritis

XX Disclosure; Page 22-24; 34pp; English.

XX This protein comprises a variant of human chitotriosidase (see  
CC also AAW31497). Its amino acid sequence was deduced from the  
CC coding sequence of a full-length cDNA clone (see AAT89181). The  
CC following are claimed: (1) a nucleic acid sequence encoding  
CC chitotriosidase; (2) a method of diagnosing a tissue remodelling  
CC disorder related to expression of a mutated chitotriosidase protein  
CC in a host comprising carrying out nucleic acid amplification; and

CC (3) a method of detecting altered expression of a chitotriosidase  
CC protein in a host comprising contacting a bodily sample with an  
CC antibody. Chitotriosidase can degrade extracellular matrix  
CC substrates with a similar carbohydrate structure to chitin. It can  
CC be used to develop products which can be used in the diagnosis and  
CC treatment of tissue remodelling disorders such as rheumatoid  
CC arthritis or atherosclerosis.

XX SQ Sequence 466 AA;

Query Match 99.7%; Score 2391; DB 18; Length 466;

Best Local Similarity 99.8%; Pred. No. 3.6e-220;

Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLCVCFYNWAGYRGEARFLPKDLDPSSLCTHLIIYAFAGMTNHQLSTTEWNETLIYQEPN 60

DB 22 AKLCVCFYNWAGYRGEARFLPKDLDPSSLCTHLIIYAFAGMTNHQLSTTEWNETLIYQEPN 81

QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 120

DB 82 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 141

QY 121 PGSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIA 180

DB 142 PGSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIA 201

QY 181 QNLDFVNLWAYDFHGSWEKVTGHSNPLYKRBESGAAASLNVDAAVQWLOKGTTPASKLI 240

DB 202 QNLDFVNLWAYDFHGSWEKVTGHSNPLYKRBESGAAASLNVDAAVQWLOKGTTPASKLI 261

QY 241 LGMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYVEVCSWGKATKORIQDOK 300

DB 262 LGMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYVEVCSWGKATKORIQDOK 321

QY 301 VPIYFRDNQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRYPLIQTL 360

DB 322 VPIYFRDNQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRYPLIQTL 381

QY 361 ROELSPLPYLPSGTPPELVKPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 420

DB 382 ROELSPLPYLPSGTPPELVKPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 441

QY 421 AGRLFQSCPTGLVFSNSCKCCTWN 445

DB 442 AGRLFQSCPTGLVFSNSCKCCTWN 466

RESULT 11

AAW31497  
ID AAW31497 standard; Protein; 464 AA.

XX AAW31497;

XX 27-APR-1998 (first entry)

XX Human chitotriosidase.

XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;  
KW rheumatoid arthritis; atherosclerosis; human.

XX Homo sapiens.

XX WO9736917-A1.

XX 09-OCT-1997.

XX 21-MAR-1997; 97WO-US05072.

XX 29-MAR-1996; 96US-0014295.

XX (HUMA-) HUMAN GENOME SCI INC.  
PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX

PI Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;  
XX WPI: 1997-503041/46.  
DR N-PSDB; AAT89180.  
XX New isolated human chitotriosidase gene - used to develop products  
PT for the diagnosis and treatment of tissue remodeling disorders, e.g.  
PT rheumatoid arthritis  
XX  
PS Disclosure; Page 19-21; 34pp; English.  
XX  
CC This protein comprises human chitotriosidase. The amino acid  
CC sequence was deduced from the coding sequence of a full-length cDNA  
CC clone (see AAT89180) deposited as ATCC 69953. A preferred variant  
CC chitotriosidase sequence is also provided (see AAW31498). The  
CC following are claimed: (1) a nucleic acid sequence encoding  
CC chitotriosidase; (2) a method of diagnosing a tissue remodeling  
CC disorder related to expression of a mutated chitotriosidase protein  
CC in a host comprising carrying out nucleic acid amplification; and  
CC (3) a method of detecting altered expression of a chitotriosidase  
CC protein in a host comprising contacting a bodily sample with an  
CC antibody. Chitotriosidase can degrade extracellular matrix  
CC substrates with a similar carbohydrate structure to chitin. It can  
CC be used to develop products which can be used in the diagnosis and  
CC treatment of tissue remodeling disorders such as rheumatoid  
CC arthritis or atherosclerosis.  
XX  
SQ Sequence 464 AA;

Query Match 98.9%; Score 2372; DB 18; Length 464;  
Best Local Similarity 99.3%; Pred. No. 2.4e-218;  
Matches 442; Conservative 0; Mismatches 1; Indels 2; Gaps 1;  
QY 1 AKLVCFYFNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTNHQSLSTTEWDETLYQBFN 60  
DB 22 AKLVCFYFNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTNHQSLSTTEWDETLYQBFN 81  
QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKSPDGLDWEY 120  
DB 82 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKSPDGLDWEY 141  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQQAQTSCKERLLLSAAVPAGQTVYDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVQDLANAFQQAQTSCKERLLLSAAVPAGQTVYDAGYEVDKIA 199  
QY 181 QNLDFFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTQPKSLI 240  
DB 200 QNLDFFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTQPKSLI 259  
QY 241 LGMPYGRSFTLLASSDTRVGAPATGSGTTPGFTKEGGMALAYEVCWKATKQRIQOK 300  
DB 260 LGMPYGRSFTLLASSDTRVGAPATGSGTTPGFTKEGGMALAYEVCWKATKQRIQOK 319  
QY 301 VPIYFRDNQWGFDDVESFKTKVSLYKQKGLGAMVWALDLDFFAGFSCNQGRYPLIOTL 360  
DB 320 VPIYFRDNQWGFDDVESFKTKVSLYKQKGLGAMVWALDLDFFAGFSCNQGRYPLIOTL 379  
QY 361 ROELSLPYLPSTGTPLEVPKQPSPEHGPSPGQDTFCOGKADGLYNPRRSSFYSKA 420  
DB 380 ROELSLPYLPSTGTPLEVPKQPSPEHGPSPGQDTFCOGKADGLYNPNRRSSFYSKA 439  
QY 421 AGRLFOQSCPTGLVFSNSCKCCTWN 445  
DB 440 AGRLFOQSCPTGLVFSNSCKCCTWN 464  
RESULT 12  
AAW40261  
ID AAW40261 standard; Protein; 373 AA.  
XX  
AC AAW40261;  
XX  
DT 15-JUN-1998 (first entry)

XX Human chitinase protein fragment.  
DE Chitinase; human; fungal infection; immunogen; diagnosis; treatment;  
XX Chitinase; human; fungal infection; immunogen; diagnosis; treatment;  
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;  
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;  
XX rheumatoid arthritis; overexpression; extracellular matrix.  
OS Homo sapiens.  
XX  
PN WO9747752-A1.  
XX  
PD 18-DEC-1997.  
XX  
PF 16-JUN-1997; 97WO-US10460.  
XX  
PR 14-JUN-1996; 96US-0663618.  
XX  
PA (ICOS-) ICOS CORP.  
XX  
PI Gray PW;  
XX  
DR WPI: 1998-052316/05.  
DR N-PSDB; AAV10435.  
XX  
PT Nucleic acids encoding human chitinase - useful as antifungal  
PT agents, especially in combination with other antifungals  
XX  
PS Claim 23; Page 48-49; 63pp; English.  
XX  
CC This sequence encodes a fragment of a novel human chitinase protein  
CC which lacks the C-terminal 72 residues of the mature protein. These  
CC residues are not critical to chitinase enzymatic activity. Chitinases are  
CC useful for treating or preventing fungal infection and as immunogens for  
CC generating antibodies which are used to purify, detect and quantify  
CC chitinases, e.g. for diagnosis of Gaucher's disease. The nucleic acid  
CC sequence of the chitinase is also useful as a probe to identify and  
CC isolate genomic DNA encoding chitinases or similar proteins, or cells  
CC expressing them or to generate transgenic ('knockout') rodents. It can  
CC also be used in hybridisation assays and to detect genetic alterations  
CC in the chitinase gene related to disease. Agents that inhibit this  
CC protein may be useful in treatment of Gaucher's disease and rheumatoid  
CC arthritis, where overexpression of the protein can damage the  
CC extracellular matrix. Chitinase also improves the activity of other  
CC antifungal agents and may allow a reduction in the dose of such agents,  
CC and thus of their side effects.  
XX  
SQ Sequence 373 AA;  
Query Match 82.7%; Score 1982; DB 19; Length 373;  
Best Local Similarity 100.0%; Pred. No. 4.3e-181;  
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKLVCFYFNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTNHQSLSTTEWDETLYQBFN 60  
DB 1 AKLVCFYFNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTNHQSLSTTEWDETLYQBFN 60  
QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKSPDGLDWEY 120  
DB 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKSPDGLDWEY 120  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQQAQTSCKERLLLSAAVPAGQTVYDAGYEVDKIA 180  
DB 121 PGSQSPAVDKERFTTLVQDLANAFQQAQTSCKERLLLSAAVPAGQTVYDAGYEVDKIA 180  
QY 181 QNLDFFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTQPKSLI 240  
DB 181 QNLDFFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTQPKSLI 240  
QY 241 LGMPYGRSFTLLASSDTRVGAPATGSGTTPGFTKEGGMALAYEVCWKATKQRIQOK 300  
DB 241 LGMPYGRSFTLLASSDTRVGAPATGSGTTPGFTKEGGMALAYEVCWKATKQRIQOK 300  
QY 301 VPIYFRDNQWGFDDVESFKTKVSLYKQKGLGAMVWALDLDFFAGFSCNQGRYPLIOTL 360

DB 301 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGAMWALDLDFAFGSCNQGRYPLIQTL 360  
 QY 361 ROELSPLPYPST 373  
 DB 361 ROELSPLPYPST 373  
 RESULT 13  
 AAY42427  
 ID AAY42427 standard; Protein; 373 AA.  
 AC AAY42427;  
 DT 10-DEC-1999 (first entry)  
 XX  
 DE Clone of the C-terminal fragment of human chitinase.  
 XX  
 KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;  
 KW organ transplant; parasite; chitin-binding; allele; vector;  
 KW truncated protein; bacterial expression.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9946390-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 12-MAR-1999; 99WO-US05343.  
 XX  
 PR 12-MAR-1998; 98US-0039198.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Gray PW, Tjoelker LW;  
 DR WPI; 1999-551417/46.  
 XX  
 PT Novel chitin-binding fragments of human chitinase used to treat fungal  
 PT infections in animals -  
 XX  
 PS Example 5; Page 66-67; 83pp; English.  
 XX  
 CC This is the amino acid sequence of the C-terminal fragment of human  
 CC chitinase, which can be expressed in bacterial cells. The fragment was  
 CC prepared by amplifying the MO-218 plasmid with primers 218-1 (AAZ21855)  
 CC and T-END (AAZ21856).  
 CC Chitinase fragments can be used to screen for proteins or other  
 CC molecules that specifically bind to the chitin-binding domain of human  
 CC chitinase or that modulate its activity. These compounds are useful for  
 CC immunization, as well as for purifying chitinase, as well as for  
 CC detection and quantification of chitinase. Polynucleotide fragments of  
 CC the invention are useful as a source of probes and primers, and to  
 CC express the proteins recombinantly. The chitinase fragments, when  
 CC conjugated to antifungal compounds, are used to treat animals,  
 CC especially humans, infected with chitin-containing parasites such as  
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
 CC sporotrichosis, and dermatophytoses.  
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
 CC protein for treating infections, especially fungal infections, is  
 CC problematic. In view of the increasing incidents of life-threatening  
 CC fungal infection in e.g. immunocompromised individuals, there exists a  
 CC need for identifying new compounds for treating fungal infection. The  
 CC chitin-binding fragments of the present invention provide this need.  
 XX  
 SQ Sequence 373 AA;  
 Query Match 82.7%; Score 1982; DB 20; Length 373;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-181;  
 Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFNWAGYRQGEARFLPKOLDPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60  
 DB 1 AKLVCFYFNWAGYRQGEARFLPKOLDPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60  
 QY 61 GLKKNPKLKTLLAIGGNWFGTKETDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120  
 DB 61 GLKKNPKLKTLLAIGGNWFGTKETDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120  
 QY 121 PGSQSPAVDKERFTTLVODLANAFQOEAQTSKGRLLLSAAVPAQOTVVDAGYEVDKIA 180  
 DB 121 PGSQSPAVDKERFTTLVODLANAFQOEAQTSKGRLLLSAAVPAQOTVVDAGYEVDKIA 180  
 QY 181 QNLDFVNLMAFYDFHGSWEKVTCHNSPLKYKROESGAASLNVDAAVQOQLXGTFPASKLI 240  
 DB 181 QNLDFVNLMAFYDFHGSWEKVTCHNSPLKYKROESGAASLNVDAAVQOQLXGTFPASKLI 240  
 QY 241 LGMPYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGLMAYEYVCSWKGATKQRIQDOK 300  
 DB 241 LGMPYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGLMAYEYVCSWKGATKQRIQDOK 300  
 QY 301 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGAMWALDLDFAFGSCNQGRYPLIQTL 360  
 DB 301 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGAMWALDLDFAFGSCNQGRYPLIQTL 360  
 QY 361 ROELSPLPYPST 373  
 DB 361 ROELSPLPYPST 373  
 RESULT 14  
 ABB76293  
 ID ABB76293 standard; Protein; 373 AA.  
 XX ABB76293;  
 AC ABB76293;  
 DT 12-AUG-2002 (first entry)  
 XX  
 DE Human chitinase truncated polypeptide.  
 XX  
 KW Chitinase; enzyme; human; fungicide; antifungal; infection;  
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;  
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
 KW Pneumocystis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6372212-B1.  
 XX  
 PD 16-APR-2002.  
 XX  
 PF 16-JUN-1997; 97US-0877599.  
 XX  
 PR 14-JUN-1996; 96US-0663618.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Gray PW;  
 DR WPI; 2002-442449/47.  
 XX  
 PT Co-administering chitinase to improve the effectiveness of fungicidal  
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -  
 XX  
 XX Example 5; Column 49-42; 26pp; English.  
 XX  
 CC The present sequence is the protein sequence of a human mature  
 CC chitinase truncated polypeptide, in which the 72 C-terminal amino  
 CC acids of the native mature protein (see ABB76291) are deleted.  
 CC The coding region for the polypeptide was obtained by PCR from  
 CC chitinase MO-218 cDNA (see ABL57380), and the polypeptide was

CC produced in transfected COS cells. This 39 kDa polypeptide lacks  
CC 6 cysteine residues of the native protein yet retains similar  
CC specific enzymatic activity. The invention provides human  
CC chitinase polynucleotides and polypeptides, and methods for the  
CC recombinant production of human chitinase products, which are  
CC expected to be useful for treating fungal infections. The  
CC chitinase can be used to improve the activity of a non-chitinase  
CC antifungal agent in the treatment of candidiasis, aspergillosis,  
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
CC histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis,  
CC mucormycosis, dermatophytoses and Pneumocystis infections.  
XX Sequence 373 AA;

Query Match 82.7%; Score 1982; DB 23; Length 373;  
Best Local Similarity 100.0%; Pred. No. 4.3e-181;  
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKLVCFYTNWAOYRQGEARFLPKDLDPSCSLTHLIYAFAGMTHQSLSTTEWNETLYQEFN 60  
DB 1 AKLVCFYTNWAOYRQGEARFLPKDLDPSCSLTHLIYAFAGMTHQSLSTTEWNETLYQEFN 60  
QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120  
DB 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180  
DB 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180  
QY 181 QNLDFVNLMAVDPHGSEWKTGNSPLRYKRBESGAAASLNVDAVQWLQKGTTPASKLI 240  
DB 181 QNLDFVNLMAVDPHGSEWKTGNSPLRYKRBESGAAASLNVDAVQWLQKGTTPASKLI 240  
QY 241 LGMPTYGRSFTLASSDTRVGAPATSGTTPGFTKEGGMLAYEYVCSWGATKQRIQDOK 300  
DB 241 LGMPTYGRSFTLASSDTRVGAPATSGTTPGFTKEGGMLAYEYVCSWGATKQRIQDOK 300  
QY 301 VPIYFRDNQWGFDDVESFETKTVSYLKQKGLGGMVWALDLDFFAGFSCNQGRIYPLIOTL 360  
DB 301 VPIYFRDNQWGFDDVESFETKTVSYLKQKGLGGMVWALDLDFFAGFSCNQGRIYPLIOTL 360  
QY 361 ROELSLPYLPSTG 373  
DB 361 ROELSLPYLPSTG 373

RESULT 15  
AAW40262  
ID AAW40262 standard; Protein; 373 AA.  
XX AC AAW40262;  
XX DT 15-JUN-1998 (first entry)  
XX DE Human chitinase protein analogue.  
XX KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;  
XX KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;  
XX KW Rheumatoid arthritis; overexpression; extracellular matrix.  
XX OS Homo sapiens.  
XX OS Synthetic.  
FH Key Location/Qualifiers  
FT Misc-difference 370 /label= P370S  
FT /note= "Wild type pro is replaced by Ser"  
XX PN W09747752-A1.  
XX PD 18-DEC-1997.

PF 16-JUN-1997; 97WO-US10460.  
XX  
PR 14-JUN-1996; 96US-0663618.  
XX  
PA (ICOS-) ICOS CORP.  
XX  
PI Gray PW;  
XX  
DR WPI; 1998-052316/05.  
XX  
PT Nucleic acids encoding human chitinase - useful as antifungal  
agents, especially in combination with other antifungals  
XX  
PS Claim 29; Page 49-50; 63pp; English.  
XX  
CC This sequence encodes a fragment of a novel human chitinase protein  
analogue in which a proline at position 370 of the wild type protein  
(see AAW40261) is substituted with a serine and the C-terminal 72  
CC residues of the mature protein are deleted. These residues are not  
CC critical to chitinase enzymatic activity. Chitinases are useful for  
CC treating or preventing fungal infection and as immunogens for generating  
CC antibodies which are used to purify, detect and quantify chitinases, e.g.  
CC for diagnosis of Gaucher's disease. The nucleic acid sequence of the  
CC chitinase is also useful as a probe to identify and isolate genomic DNA  
CC encoding chitinases or similar proteins, or cells expressing them or to  
CC generate transgenic ('knockout') rodents. It can also be used in  
CC hybridisation assays and to detect genetic alterations in the chitinase  
CC gene related to disease. Agents that inhibit this protein may be useful  
CC in treatment of Gaucher's disease and rheumatoid arthritis, where  
CC overexpression of the protein can damage the extracellular matrix.  
CC Chitinase also improves the activity of other antifungal agents and may  
CC allow a reduction in the dose of such agents, and thus of their side  
CC effects.  
XX

SQ Sequence 373 AA;  
Query Match 82.3%; Score 1974; DB 19; Length 373;  
Best Local Similarity 99.7%; Pred. No. 2.5e-180;  
Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AKLVCFYTNWAOYRQGEARFLPKDLDPSCSLTHLIYAFAGMTHQSLSTTEWNETLYQEFN 60  
DB 1 AKLVCFYTNWAOYRQGEARFLPKDLDPSCSLTHLIYAFAGMTHQSLSTTEWNETLYQEFN 60  
QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120  
DB 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180  
DB 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180  
QY 181 QNLDFVNLMAVDPHGSEWKTGNSPLRYKRBESGAAASLNVDAVQWLQKGTTPASKLI 240  
DB 181 QNLDFVNLMAVDPHGSEWKTGNSPLRYKRBESGAAASLNVDAVQWLQKGTTPASKLI 240  
QY 241 LGMPTYGRSFTLASSDTRVGAPATSGTTPGFTKEGGMLAYEYVCSWGATKQRIQDOK 300  
DB 241 LGMPTYGRSFTLASSDTRVGAPATSGTTPGFTKEGGMLAYEYVCSWGATKQRIQDOK 300  
QY 301 VPIYFRDNQWGFDDVESFETKTVSYLKQKGLGGMVWALDLDFFAGFSCNQGRIYPLIOTL 360  
DB 301 VPIYFRDNQWGFDDVESFETKTVSYLKQKGLGGMVWALDLDFFAGFSCNQGRIYPLIOTL 360  
QY 361 ROELSLPYLPSTG 373  
DB 361 ROELSLPYLPSTG 373

Search completed: June 29, 2003, 20:59:40  
Job time : 30.6228 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:57:59 ; Search time 11.52 Seconds  
(without alignments)  
1136.565 Million cell updates/sec

Title: US-10-004-219b-10  
Perfect score: 2398  
Sequence: 1 AKLVCFYTNWAQYRQGEARF.....QQSCPTGLVFSNCKCCTWN 445

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/6CTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	100.0	466	2	US-08-486-839-4
2	2398	100.0	466	3	US-09-151-011-4
3	2398	100.0	466	4	US-09-039-198A-2
4	2398	100.0	466	4	US-09-343-623-4
5	2398	100.0	466	4	US-08-877-599-2
6	2398	100.0	466	4	US-09-267-574-2
7	2392	99.7	466	4	US-09-039-198A-4
8	2392	99.7	466	4	US-08-877-599-4
9	2392	99.7	466	4	US-09-267-574-4
10	1982	82.7	373	4	US-09-039-198A-14
11	1982	82.7	373	4	US-08-877-599-14
12	1982	82.7	373	4	US-09-267-574-14
13	1974	82.3	373	4	US-09-039-198A-15
14	1974	82.3	373	4	US-08-877-599-15
15	1974	82.3	373	4	US-09-267-574-15
16	1935	80.7	387	2	US-08-486-839-6
17	1935	80.7	387	3	US-09-151-011-6
18	1935	80.7	387	4	US-09-343-623-6
19	1045.5	43.6	385	2	US-08-694-915-2
20	1045.5	43.6	416	2	US-08-694-915-4
21	1021.5	42.6	383	4	US-09-459-749D-17
22	824	34.4	554	3	US-08-524-051-2
23	824	34.4	554	3	US-09-052-778-16
24	803.5	33.5	559	4	US-09-545-814-14
25	803.5	33.5	583	4	US-09-545-814-2
26	803.5	33.5	583	4	US-09-545-814-5
27	797.5	33.3	635	4	US-09-545-814-32

28	762	31.8	490	4	US-09-292-225-41	Sequence 41, Appl
29	762	31.8	509	4	US-09-292-225-35	Sequence 35, Appl
30	762	31.8	509	4	US-09-292-225-38	Sequence 38, Appl
31	738	30.8	536	4	US-09-292-225-21	Sequence 21, Appl
32	738	30.8	555	4	US-09-292-225-15	Sequence 15, Appl
33	738	30.8	555	4	US-09-292-225-18	Sequence 18, Appl
34	660	27.5	489	4	US-09-545-814-29	Sequence 29, Appl
35	499.5	20.8	389	1	US-07-939-501A-1	Sequence 1, Appl
36	499.5	20.8	389	4	US-08-448-398-7	Sequence 7, Appl
37	499.5	20.8	423	1	US-07-939-501A-10	Sequence 10, Appl
38	499.5	20.8	423	1	US-07-939-501A-12	Sequence 12, Appl
39	486	20.3	424	1	US-08-045-269C-2	Sequence 2, Appl
40	486	20.3	424	3	US-08-371-680-2	Sequence 2, Appl
41	485	20.2	424	5	PCT-US94-01198-2	Sequence 2, Appl
42	474	19.8	442	3	US-09-052-778-2	Sequence 2, Appl
43	462.5	19.3	377	2	US-08-591-629-8	Sequence 8, Appl
44	460.5	19.2	371	2	US-08-591-629-2	Sequence 2, Appl
45	444.5	18.5	440	3	US-09-052-778-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1  
US-08-486-839-4  
; Sequence 4, Application US/08486839  
; Patent No. 5928928  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A human chitinase, its recombinant  
; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11758  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,839  
; FILING DATE: 07 - June - 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baron, Ronald J.  
; REGISTRATION NUMBER: 29,281  
; REFERENCE/DOCKET NUMBER: 294-26  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
US-08-486-839-4

Query Match 100.0%; Score 2398; DB 2; Length 466;  
Best Local Similarity 100.0%; Pred. No. 5,1e-218;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAQYRQGEARFLPKDLPSCLTGLIYAFAGTWHQSTTWNDETLYQEFN 60  
DB 22 AKLVCFYTNWAQYRQGEARFLPKDLPSCLTGLIYAFAGTWHQSTTWNDETLYQEFN 81

QY 61 GLKKNPKLKTLLAIGGNFTQKFTDMVATANNROTFFVNSAIRFLRKYSFQDGLDLIDWEY 120  
DB 82 GLKKNPKLKTLLAIGGNFTQKFTDMVATANNROTFFVNSAIRFLRKYSFQDGLDLIDWEY 141  
QY 121 PGSGSPAVDKERFTTLVQDLANAFQOEATSGKERLLLSAAVPAGQTYVDAGYVDKIA 180  
DB 142 PGSGSPAVDKERFTTLVQDLANAFQOEATSGKERLLLSAAVPAGQTYVDAGYVDKIA 201  
QY 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVQOVLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVQOVLQKGTTPASKLI 261  
QY 241 LGNPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 300  
DB 262 LGNPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 321  
QY 301 VPYIFRDNQWGFDDVESFRTKVSYLKQKGLGAMVWALDLDFFAGFSCNQGRIPLIOTL 360  
DB 322 VPYIFRDNQWGFDDVESFRTKVSYLKQKGLGAMVWALDLDFFAGFSCNQGRIPLIOTL 381  
QY 361 ROELSPLYPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRSSSFYSCA 420  
DB 382 ROELSPLYPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRSSSFYSCA 441  
QY 421 AGRLFQOQSCPTGLVFSNSCKCCTWN 445  
DB 442 AGRLFQOQSCPTGLVFSNSCKCCTWN 466

## RESULT 2

US-09-151-011-4  
; Sequence 4, Application US/09151011  
; Patent No. 6057142  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant  
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in  
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 6900 Jericho Turnpike  
; CITY: Syosset  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11791  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; SOFTWARE:

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/151,011  
; FILING DATE: 10 - September - 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Morris, Robert C.  
; REGISTRATION NUMBER: 42,910  
; REFERENCE/DOCKET NUMBER: 294-32 DIV  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-09-151-011-4

Query Match 100.0%; Score 2398; DB 3; Length 466;  
Best Local Similarity 100.0%; Pred. No. 5,1e-218;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKLVCFYTNWAOYRQGEARFLPKDLPSLCTHLIYAFAGMTNHLSTTEWDETLYOEFN 60  
DB 22 AKLVCFYTNWAOYRQGEARFLPKDLPSLCTHLIYAFAGMTNHLSTTEWDETLYOEFN 81  
QY 61 GLKKNPKLKTLLAIGGNFTQKFTDMVATANNROTFFVNSAIRFLRKYSFQDGLDLIDWEY 120  
DB 82 GLKKNPKLKTLLAIGGNFTQKFTDMVATANNROTFFVNSAIRFLRKYSFQDGLDLIDWEY 141  
QY 121 PGSGSPAVDKERFTTLVQDLANAFQOEATSGKERLLLSAAVPAGQTYVDAGYVDKIA 180  
DB 142 PGSGSPAVDKERFTTLVQDLANAFQOEATSGKERLLLSAAVPAGQTYVDAGYVDKIA 201  
QY 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVQOVLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVQOVLQKGTTPASKLI 261  
QY 241 LGNPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 300  
DB 262 LGNPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 321  
QY 301 VPYIFRDNQWGFDDVESFRTKVSYLKQKGLGAMVWALDLDFFAGFSCNQGRIPLIOTL 360  
DB 322 VPYIFRDNQWGFDDVESFRTKVSYLKQKGLGAMVWALDLDFFAGFSCNQGRIPLIOTL 381  
QY 361 ROELSPLYPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRSSSFYSCA 420  
DB 382 ROELSPLYPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRSSSFYSCA 441  
QY 421 AGRLFQOQSCPTGLVFSNSCKCCTWN 445  
DB 442 AGRLFQOQSCPTGLVFSNSCKCCTWN 466

## RESULT 3

US-09-039-198A-2  
; Sequence 2, Application US/09039198A  
; Patent No. 6200951  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,198A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/34391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid



TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-039-198A-2

Query Match 100.0%; Score 2398; DB 4; Length 466;  
Best Local Similarity 100.0%; Pred. No. 5.1e-218;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLDPSCSLTHLIYAFAGMTNHQSLSTTEWDETLYQEFN 60  
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLDPSCSLTHLIYAFAGMTNHQSLSTTEWDETLYQEFN 81  
QY 61 GLKKNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 120  
DB 82 GLKKNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 141  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQQAOTSGKERLLLSAAVPAQOTYVDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVQDLANAFQQAOTSGKERLLLSAAVPAQOTYVDAGYEVDKIA 201  
QY 181 QNLDFVNLMAVDPHGSEWEKVTGHNSPLYKROESGAAASLNDAVQVWLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDPHGSEWEKVTGHNSPLYKROESGAAASLNDAVQVWLQKGTTPASKLI 261  
QY 241 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKATKQRIQDOK 300  
DB 262 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKATKQRIQDOK 321  
QY 301 VPIYFRDQWGVDFDVESEFKTKVSKYKOKGLGGAMVWALDDDFAGFCNQGRIYPLIOTL 360  
DB 322 VPIYFRDQWGVDFDVESEFKTKVSKYKOKGLGGAMVWALDDDFAGFCNQGRIYPLIOTL 381  
QY 361 ROELSLPVLPSGTPELEVPKQPSPEHGPSPGQDTFCQKADGLYNPNRERSFYSCA 420  
DB 382 ROELSLPVLPSGTPELEVPKQPSPEHGPSPGQDTFCQKADGLYNPNRERSFYSCA 441  
QY 421 AGRLFQSCPTGLVFSNCKCCTWN 445  
DB 442 AGRLFQSCPTGLVFSNCKCCTWN 466

RESULT 4

US-09-343-623-4  
; Sequence 4, Application US/09343623  
; Patent No. 6303118  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A human chitinase, its recombinant  
; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11758  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/343,623  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,839  
; FILING DATE: 07-June-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baron, Ronald J.  
; REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 466 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-09-343-623-4

Query Match 100.0%; Score 2398; DB 4; Length 466;

Best Local Similarity 100.0%; Pred. No. 5.1e-218;

Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLDPSCSLTHLIYAFAGMTNHQSLSTTEWDETLYQEFN 60  
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLDPSCSLTHLIYAFAGMTNHQSLSTTEWDETLYQEFN 81  
QY 61 GLKKNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 120  
DB 82 GLKKNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 141  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQQAOTSGKERLLLSAAVPAQOTYVDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVQDLANAFQQAOTSGKERLLLSAAVPAQOTYVDAGYEVDKIA 201  
QY 181 QNLDFVNLMAVDPHGSEWEKVTGHNSPLYKROESGAAASLNDAVQVWLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDPHGSEWEKVTGHNSPLYKROESGAAASLNDAVQVWLQKGTTPASKLI 261  
QY 241 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKATKQRIQDOK 300  
DB 262 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKATKQRIQDOK 321  
QY 301 VPIYFRDQWGVDFDVESEFKTKVSKYKOKGLGGAMVWALDDDFAGFCNQGRIYPLIOTL 360  
DB 322 VPIYFRDQWGVDFDVESEFKTKVSKYKOKGLGGAMVWALDDDFAGFCNQGRIYPLIOTL 381  
QY 361 ROELSLPVLPSGTPELEVPKQPSPEHGPSPGQDTFCQKADGLYNPNRERSFYSCA 420  
DB 382 ROELSLPVLPSGTPELEVPKQPSPEHGPSPGQDTFCQKADGLYNPNRERSFYSCA 441  
QY 421 AGRLFQSCPTGLVFSNCKCCTWN 445  
DB 442 AGRLFQSCPTGLVFSNCKCCTWN 466

RESULT 5

US-08-877-599-2

; Sequence 2, Application US/08877599

; Patent No. 6372212

; GENERAL INFORMATION:

; APPLICANT: Gray, Patrick W.

; TITLE OF INVENTION: Chitinase Materials and Methods

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/877,599

/ FILING DATE: 514  
/ CLASSIFICATION: 514  
/ PRIOR APPLICATION DATA: US 08/663,618  
/ APPLICATION NUMBER: 514  
/ FILING DATE: 14-JUN-1996  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Rin-Laures, Li-Hsien  
/ REGISTRATION NUMBER: 33,547  
/ REFERENCE/DOCKET NUMBER: 27866/33994  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 312/474-6300  
/ TELEFAX: 312/474-0448  
/ TELEX: 25-3856  
/ INFORMATION FOR SEQ ID NO: 2:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 466 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ US-08-877-599-2

Query Match 100.0%; Score 2398; DB 4; Length 466;  
Best Local Similarity 100.0%; Pred. No. 5.1e-218;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNNHQLSTTWNDETLYQEFN 60  
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNNHQLSTTWNDETLYQEFN 81  
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPFGDLDDWEY 120  
DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPFGDLDDWEY 141  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 201  
QY 181 QNLDFVNLMAFYDGHSGWEKVTGHNPSPLYKROESGAAASLNVDAVQOVLKGTTPASKLI 240  
DB 202 QNLDFVNLMAFYDGHSGWEKVTGHNPSPLYKROESGAAASLNVDAVQOVLKGTTPASKLI 261  
QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWGKATKQRIQDOK 300  
DB 262 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWGKATKQRIQDOK 321  
QY 301 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 360  
DB 322 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 381  
QY 361 ROELSPLPSPGTPPELVPKQPSPEHGPSPGQDTFCQKADGLYPNPERSSFYSICA 420  
DB 382 ROELSPLPSPGTPPELVPKQPSPEHGPSPGQDTFCQKADGLYPNPERSSFYSICA 441  
QY 421 AGRLFQOCSPTGLVFSNCKCCTWN 445  
DB 442 AGRLFQOCSPTGLVFSNCKCCTWN 466

RESULT 6  
US-09-267-574-2  
Sequence 2, Application US/09267574  
Patent No. 6399571  
GENERAL INFORMATION:  
APPLICANT: Tjoelker, Larry W.  
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
FILE REFERENCE: 27866/35407  
CURRENT APPLICATION NUMBER: US/09/267,574  
EARLIER FILING DATE: 1999-03-12  
EARLIER APPLICATION NUMBER: 09/039,198  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: Patent in Ver. 2.0

/ SEQ ID NO 2  
/ LENGTH: 466  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
/ US-09-267-574-2

Query Match 100.0%; Score 2398; DB 4; Length 466;  
Best Local Similarity 100.0%; Pred. No. 5.1e-218;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNNHQLSTTWNDETLYQEFN 60  
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNNHQLSTTWNDETLYQEFN 81  
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPFGDLDDWEY 120  
DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPFGDLDDWEY 141  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 201  
QY 181 QNLDFVNLMAFYDGHSGWEKVTGHNPSPLYKROESGAAASLNVDAVQOVLKGTTPASKLI 240  
DB 202 QNLDFVNLMAFYDGHSGWEKVTGHNPSPLYKROESGAAASLNVDAVQOVLKGTTPASKLI 261  
QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWGKATKQRIQDOK 300  
DB 262 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWGKATKQRIQDOK 321  
QY 301 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 360  
DB 322 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 381  
QY 361 ROELSPLPSPGTPPELVPKQPSPEHGPSPGQDTFCQKADGLYPNPERSSFYSICA 420  
DB 382 ROELSPLPSPGTPPELVPKQPSPEHGPSPGQDTFCQKADGLYPNPERSSFYSICA 441  
QY 421 AGRLFQOCSPTGLVFSNCKCCTWN 445  
DB 442 AGRLFQOCSPTGLVFSNCKCCTWN 466

RESULT 7  
US-09-039-198A-4  
Sequence 4, Application US/09039198A  
Patent No. 6200951  
GENERAL INFORMATION:  
APPLICANT: Gray, Patrick W.  
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,198A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 27866/34391  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-039-198A-4

Query Match 99.7%; Score 2392; DB 4; Length 466;  
Best Local Similarity 99.8%; Pred. No. 1.9e-217;  
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAOYRQGEARFLPKDLPDLPSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 60  
DB 22 AKLVCFYTNWAOYRQGEARFLPKDLPDLPSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 81  
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120  
DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 141  
QY 121 PGSQSPAVDKERFTTLVODLANAQOEATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVODLANAQOEATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 201  
QY 181 QNLDFVNLMAVDFHGSWEKVTGHNPSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDFHGSWEKVTGHNPSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 261  
QY 241 LCMPTYGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGLMAYEVCWKGATKQRIQDOK 300  
DB 262 LCMPTYGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGLMAYEVCWKGATKQRIQDOK 321  
QY 301 VPIYFRDNQWGFDDVESFKTKVYLKQGLGAMVWALDDDFAGFSCNQGRYPLIQT 360  
DB 322 VPIYFRDNQWGFDDVESFKTKVYLKQGLGAMVWALDDDFAGFSCNQGRYPLIQT 381  
QY 361 ROELSLPYLPSTGTPLEVPKQPEPEHGPSPGQDTFCQKADGLYNPNRPRSSFYSCA 420  
DB 382 ROELSLPYLPSTGTPLEVPKQPEPEHGPSPGQDTFCQKADGLYNPNRPRSSFYSCA 441  
QY 421 AGRLFQOQSCPTGLVFSNSCKCTWN 445  
DB 442 AGRLFQOQSCPTGLVFSNSCKCTWN 466

## RESULT 8

US-08-877-599-4

Sequence 4, Application US/08877599

Patent No. 6372212

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

TITLE OF INVENTION: Chitinase Materials and Methods

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/877,599

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/663,618

FILING DATE: 14-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 27866/33994  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 466 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-877-599-4

Query Match 99.7%; Score 2392; DB 4; Length 466;

Best Local Similarity 99.8%; Pred. No. 1.9e-217;

Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAOYRQGEARFLPKDLPDLPSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 60  
DB 22 AKLVCFYTNWAOYRQGEARFLPKDLPDLPSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 81  
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120  
DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 141  
QY 121 PGSQSPAVDKERFTTLVODLANAQOEATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVODLANAQOEATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 201  
QY 181 QNLDFVNLMAVDFHGSWEKVTGHNPSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDFHGSWEKVTGHNPSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 261  
QY 241 LCMPTYGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGLMAYEVCWKGATKQRIQDOK 300  
DB 262 LCMPTYGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGLMAYEVCWKGATKQRIQDOK 321  
QY 301 VPIYFRDNQWGFDDVESFKTKVYLKQGLGAMVWALDDDFAGFSCNQGRYPLIQT 360  
DB 322 VPIYFRDNQWGFDDVESFKTKVYLKQGLGAMVWALDDDFAGFSCNQGRYPLIQT 381  
QY 361 ROELSLPYLPSTGTPLEVPKQPEPEHGPSPGQDTFCQKADGLYNPNRPRSSFYSCA 420  
DB 382 ROELSLPYLPSTGTPLEVPKQPEPEHGPSPGQDTFCQKADGLYNPNRPRSSFYSCA 441  
QY 421 AGRLFQOQSCPTGLVFSNSCKCTWN 445  
DB 442 AGRLFQOQSCPTGLVFSNSCKCTWN 466

## RESULT 9

US-09-267-574-4

Sequence 4, Application US/09267574

Patent No. 6399571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

APPLICANT: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

FILE REFERENCE: 27866/35407

CURRENT APPLICATION NUMBER: US/09/267,574

CURRENT FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: 09/039,198

EARLIER FILING DATE: 1998-03-12

NUMBER OF SEQ ID NOS: 39

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 4

LENGTH: 466

TYPE: PRT

ORGANISM: Homo sapiens

US-09-267-574-4

Query Match 99.7%; Score 2392; DB 4; Length 466;  
Best Local Similarity 99.8%; Pred. No. 1.9e-217;  
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYQEFN 60  
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYQEFN 81  
QY 61 GLKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTTFVNSAIRFLRKYSPDGLDLDWEY 120  
DB 82 GLKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTTFVNSAIRFLRKYSPDGLDLDWEY 141  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 180  
DB 142 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 201  
QY 181 QNLDFVNLMAVDPHGSEKVTGHNSPLYKROESGAAASLNVDAAVQOVLQKGTTPASKLI 240  
DB 202 QNLDFVNLMAVDPHGSEKVTGHNSPLYKROESGAAASLNVDAAVQOVLQKGTTPASKLI 261  
QY 241 LGMPTTGRSFTLASSSDTRVGAPATSGTGPFTKEGMLAYEVCWSKGATKQRIQDOK 300  
DB 262 LGMPTTGRSFTLASSSDTRVGAPATSGTGPFTKEGMLAYEVCWSKGATKQRIQDOK 321  
QY 301 VPIYFRDNQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIYPLIOTL 360  
DB 322 VPIYFRDNQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIYPLIOTL 381  
QY 361 ROELSPLYPSTGTELEVPKPGQSPBEPHGPSQDTPFCQKADGLYPNPRSSFYSCA 420  
DB 382 ROELSPLYPSTGTELEVPKPGQSPBEPHGPSQDTPFCQKADGLYPNPRSSFYSCA 441  
QY 421 AGRLFQSCPTGLVFNSSCKCTWN 445  
DB 442 AGRLFQSCPTGLVFNSSCKCTWN 466

RESULT 10

US-09-039-198A-14  
; Sequence 14, Application US/09039198A  
; Patent No. 6200951  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Toelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,198A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/34391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLSCULE TYPE: peptide  
US-09-039-198A-14

Query Match 82.7%; Score 1982; DB 4; Length 373;

Best Local Similarity 100.0%; Pred. No. 7.1e-179;  
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYQEFN 60  
DB 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYQEFN 60  
QY 61 GLKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTTFVNSAIRFLRKYSPDGLDLDWEY 120  
DB 61 GLKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTTFVNSAIRFLRKYSPDGLDLDWEY 120  
QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 180  
DB 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 180  
QY 181 QNLDFVNLMAVDPHGSEKVTGHNSPLYKROESGAAASLNVDAAVQOVLQKGTTPASKLI 240  
DB 181 QNLDFVNLMAVDPHGSEKVTGHNSPLYKROESGAAASLNVDAAVQOVLQKGTTPASKLI 240  
QY 241 LGMPTTGRSFTLASSSDTRVGAPATSGTGPFTKEGMLAYEVCWSKGATKQRIQDOK 300  
DB 241 LGMPTTGRSFTLASSSDTRVGAPATSGTGPFTKEGMLAYEVCWSKGATKQRIQDOK 300  
QY 301 VPIYFRDNQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIYPLIOTL 360  
DB 301 VPIYFRDNQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIYPLIOTL 360  
QY 361 ROELSPLYPSTG 373  
DB 361 ROELSPLYPSTG 373

RESULT 11

US-08-877-599-14  
; Sequence 14, Application US/08877599  
; Patent No. 6372212  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; TITLE OF INVENTION: Chitinase Materials and Methods  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/877,599  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/663,618  
; FILING DATE: 14-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/33994  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-877-599-14

Query Match 82.7%; Score 1982; DB 4; Length 373;

Best Local Similarity 100.0%; Pred. No. 7.1e-179;

Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60

DB 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60

QY 61 GLKXNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120

DB 61 GLKXNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120

QY 121 PGSQSPAVDKERFTTLVODLANAQOEATSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180

DB 121 PGSQSPAVDKERFTTLVODLANAQOEATSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180

QY 181 QNLDFVNLWAYDFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 240

DB 181 QNLDFVNLWAYDFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 240

QY 241 LGMPYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGLAYEYVCSWKGATKQRIQDOK 300

DB 241 LGMPYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGLAYEYVCSWKGATKQRIQDOK 300

QY 301 VPIYFRDNQWGFDDVESFKTKVSLKQGLGGMVWALDLDLDFAGFSCNOGRYPLIOTL 360

DB 301 VPIYFRDNQWGFDDVESFKTKVSLKQGLGGMVWALDLDLDFAGFSCNOGRYPLIOTL 360

QY 361 ROELSLPYLPSTGT 373

DB 361 ROELSLPYLPSTGT 373

RESULT 12

US-09-267-574-14

Sequence 14, Application US/09267574

Patent No. 6399571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

FILE REFERENCE: 27866/35407

CURRENT APPLICATION NUMBER: US/09/267,574

CURRENT FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: 09/039,198

EARLIER FILING DATE: 1998-03-12

NUMBER OF SEQ ID NOS: 39

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 14

LENGTH: 373

TYPE: PRT

ORGANISM: Homo sapiens

US-09-267-574-14

Query Match

Best Local Similarity 82.7%; Score 1982; DB 4; Length 373;

Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60

DB 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60

QY 61 GLKXNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120

DB 61 GLKXNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120

QY 121 PGSQSPAVDKERFTTLVODLANAQOEATSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180

DB 121 PGSQSPAVDKERFTTLVODLANAQOEATSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180

QY 181 QNLDFVNLWAYDFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 240

DB 181 QNLDFVNLWAYDFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 240

QY 241 LGMPYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGLAYEYVCSWKGATKQRIQDOK 300

DB 241 LGMPYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGLAYEYVCSWKGATKQRIQDOK 300

QY 301 VPIYFRDNQWGFDDVESFKTKVSLKQGLGGMVWALDLDLDFAGFSCNOGRYPLIOTL 360

DB 301 VPIYFRDNQWGFDDVESFKTKVSLKQGLGGMVWALDLDLDFAGFSCNOGRYPLIOTL 360

QY 361 ROELSLPYLPSTGT 373

DB 361 ROELSLPYLPSTGT 373

RESULT 13

US-09-039-198A-15

Sequence 15, Application US/09039198A

Patent No. 6200951

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

APPLICANT: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray &amp; Borun

STREET: 233 South Wacker Drive/6300 Sears Tower

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/039,198A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien

REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 27866/34391

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 474-6300

TELEFAX: (312) 474-0448

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-039-198A-15

Query Match

Best Local Similarity 82.3%; Score 1974; DB 4; Length 373;

Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60

DB 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60

QY 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEY 120  
DB 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEY 120  
QY 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180  
DB 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180  
QY 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLKGTTPASKLI 240  
DB 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLKGTTPASKLI 240  
QY 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEYVCSWKATKQRIQDOK 300  
DB 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEYVCSWKATKQRIQDOK 300  
QY 301 VPYIFRDNOMVGFDDVESFRTKVSYLKQKGLGAMWALDLDLDFAGFSCNQGRYPLIOTL 360  
DB 301 VPYIFRDNOMVGFDDVESFRTKVSYLKQKGLGAMWALDLDLDFAGFSCNQGRYPLIOTL 360  
QY 361 ROELSPLPLPSGT 373  
DB 361 ROELSPLPLSSGT 373

## RESULT 14

US-08-877-599-15  
; Sequence 15, Application US/08877599  
; Patent No. 6372212  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; TITLE OF INVENTION: Chitinase Materials and Methods  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/877,599  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/663,618  
; FILING DATE: 14-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/33994  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 373 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-877-599-15

Query Match 82.3%; Score 1974; DB 4; Length 373;  
Best Local Similarity 99.7%; Pred. No. 4e-178;  
Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAQYRQGEARFLPKDLDPSCSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60  
DB 1 AKLVCFYTNWAQYRQGEARFLPKDLDPSCSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60  
QY 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEY 120  
DB 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEY 120  
QY 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180  
DB 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180  
QY 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLKGTTPASKLI 240  
DB 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLKGTTPASKLI 240  
QY 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEYVCSWKATKQRIQDOK 300  
DB 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEYVCSWKATKQRIQDOK 300  
QY 301 VPYIFRDNOMVGFDDVESFRTKVSYLKQKGLGAMWALDLDLDFAGFSCNQGRYPLIOTL 360  
DB 301 VPYIFRDNOMVGFDDVESFRTKVSYLKQKGLGAMWALDLDLDFAGFSCNQGRYPLIOTL 360  
QY 361 ROELSPLPLPSGT 373  
DB 361 ROELSPLPLSSGT 373

## RESULT 15

US-09-267-574-15  
; Sequence 15, Application US/09267574  
; Patent No. 6399571  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; TITLE OF INVENTION: Chitinase  
; FILE REFERENCE: 27866/35407  
; CURRENT APPLICATION NUMBER: US/09/267,574  
; CURRENT FILING DATE: 1999-03-12  
; EARLIER APPLICATION NUMBER: 09/039,198  
; EARLIER FILING DATE: 1998-03-12  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-267-574-15

Query Match 82.3%; Score 1974; DB 4; Length 373;  
Best Local Similarity 99.7%; Pred. No. 4e-178;  
Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAQYRQGEARFLPKDLDPSCSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60  
DB 1 AKLVCFYTNWAQYRQGEARFLPKDLDPSCSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60  
QY 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEY 120  
DB 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEY 120  
QY 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180  
DB 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180  
QY 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLKGTTPASKLI 240  
DB 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLKGTTPASKLI 240  
QY 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEYVCSWKATKQRIQDOK 300

Db 241 LCMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYYEVCSWKATKQRIQDK 300  
Qy 301 VFYIFRDNQVGFDDVESEFKTKVSYLKQKGLGAMWALDLDLDDFAGFSCNQGRYPLIOTL 360  
Db 301 VFYIFRDNQVGFDDVESEFKTKVSYLKQKGLGAMWALDLDLDDFAGFSCNQGRYPLIOTL 360  
Qy 361 ROELSLPYLPSGT 373  
Db 361 ROELSLPYLSSGT 373

Search completed: June 29, 2003, 21:03:29  
Job time : 12.52 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 21:01:56 ; Search time 17.1154 Seconds  
(without alignments)  
2850.818 Million cell updates/sec

Title: US-10-004-219B-10  
Perfect score: 2398  
Sequence: 1 AKLVCFYTNWAGYRGEARF.....QQSCPTGLVFNCKCTWN 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	100.0	445	9	US-10-004-219B-10
2	1323	55.2	455	9	US-10-004-219B-14
3	1323	55.2	476	9	US-10-004-219B-1
4	1301.5	54.3	452	9	US-10-004-219B-9
5	1301.5	54.3	473	9	US-10-004-219B-4
6	1041.5	43.4	383	9	US-10-097-340-45
7	1021.5	42.6	383	10	US-09-459-749D-17
8	762	31.8	490	9	US-10-218-743-41
9	762	31.8	509	9	US-10-218-743-35
10	762	31.8	509	9	US-10-218-743-38
11	738	30.8	536	9	US-10-218-743-21
12	738	30.8	555	9	US-10-218-743-15
13	738	30.8	555	9	US-10-218-743-18
14	487	20.3	371	9	US-09-923-844B-2
15	335.5	14.0	376	10	US-09-748-033-3
16	214	8.9	170	9	US-10-218-743-44
17	101.5	4.2	1289	10	US-09-738-363-4
18	101	4.2	411	9	US-09-866-050A-302
19	100	4.2	802	10	US-09-757-049A-1

20	97.5	4.1	346	9	US-10-260-877-80	Sequence 80, Appl
21	97	4.0	393	9	US-09-977-418-12	Sequence 12, Appl
22	97	4.0	393	9	US-09-977-033A-12	Sequence 12, Appl
23	97	4.0	393	10	US-09-840-787-11	Sequence 11, Appl
24	96.5	4.0	1276	10	US-09-982-610-24	Sequence 24, Appl
25	96	4.0	416	9	US-10-023-282-461	Sequence 461, Appl
26	94	3.9	289	9	US-09-957-187-14	Sequence 14, Appl
27	94	3.9	1154	9	US-10-224-624-8	Sequence 8, Appl
28	92.5	3.9	298	10	US-09-748-033-1	Sequence 1, Appl
29	92	3.8	363	9	US-10-023-282-278	Sequence 278, Appl
30	91.5	3.8	1140	9	US-09-974-973-19	Sequence 19, Appl
31	91.5	3.8	1140	9	US-09-738-626-4265	Sequence 4265, Appl
32	91.5	3.8	1140	9	US-10-045-072-2	Sequence 2, Appl
33	91.5	3.8	1157	9	US-09-974-973-2	Sequence 4, Appl
34	91.5	3.8	1157	9	US-09-974-973-4	Sequence 4, Appl
35	91.5	3.8	1190	9	US-09-964-295-2	Sequence 2, Appl
36	91.5	3.8	2322	9	US-09-919-039-15	Sequence 15, Appl
37	91	3.8	408	9	US-10-027-806-74	Sequence 74, Appl
38	91	3.8	408	9	US-10-034-623-74	Sequence 74, Appl
39	91	3.8	408	9	US-10-027-801-74	Sequence 74, Appl
40	91	3.8	1293	10	US-09-815-242-10079	Sequence 10079, A
41	90.5	3.8	2783	10	US-09-816-669A-14	Sequence 14, Appl
42	90	3.8	2971	9	US-10-146-473-50	Sequence 50, Appl
43	89.5	3.7	222	9	US-09-727-855B-5	Sequence 5, Appl
44	89.5	3.7	1079	9	US-10-112-488-39	Sequence 39, Appl
45	89	3.7	235	9	US-09-738-626-5245	Sequence 5245, Appl

ALIGNMENTS

RESULT 1  
US-10-004-219B-10  
; Sequence 10, Application US/10004219B  
; Publication No. US2003008741A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-513605  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 445  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)..(445)  
; OTHER INFORMATION: /notes="Human chitotriosidase"  
US-10-004-219B-10

Query Match 100.0%; Score 2398; DB 9; Length 445;  
Best Local Similarity 100.0%; Pred. No. 5.2e-202;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AKLVCFYTNWAGYRGEARFLPKDLPDLCTHLIYAFAGMTNHLSTTEWNETLYQEFN 60
DB	1	AKLVCFYTNWAGYRGEARFLPKDLPDLCTHLIYAFAGMTNHLSTTEWNETLYQEFN 60
QY	61	GLKXGNPKLTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLDWEY 120
DB	61	GLKXGNPKLTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLDWEY 120
QY	121	PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAMPAGQTVVDAGYEVDKIA 180
DB	121	PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAMPAGQTVVDAGYEVDKIA 180

181 QNLDVNLMAVDYDHFSGMEKVTGHSNPLYKQREBSGAASLNVDAAVQOVLKQGTTPASKLI 240  
181 QNLDVNLMAVDYDHFSGMEKVTGHSNPLYKQREBSGAASLNVDAAVQOVLKQGTTPASKLI 240  
241 LGMPTVGRSFTLASSSDTRVGAPATGSGTPGPTKGGMLAYEVCWSKGATKQRIODQK 300  
241 LGMPTVGRSFTLASSSDTRVGAPATGSGTPGPTKGGMLAYEVCWSKGATKQRIODQK 300  
301 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGGAMVWALDLDLDDFAGFSCNOGRYPILOT 360  
301 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGGAMVWALDLDLDDFAGFSCNOGRYPILOT 360  
361 ROELSPLPSPGTPPELEVPRKPGQSPPEHGPSPGQDTFCQKADGLYPNPRSSFSYSCA 420  
361 ROELSPLPSPGTPPELEVPRKPGQSPPEHGPSPGQDTFCQKADGLYPNPRSSFSYSCA 420  
421 AGRLFQOCSCTGLVFNNSCKCCTWN 445  
421 AGRLFQOCSCTGLVFNNSCKCCTWN 445

RESULT 2  
US-10-004-219B-14  
; Sequence 14, Application US/10004219B  
; Publication No. US20030087414A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: (1)..(455)  
; OTHER INFORMATION: /note="Human AMCase"  
US-10-004-219B-14

Query Match 55.2%; Score 1323; DB 9; Length 455;  
Best Local Similarity 52.2%; Pred. No. 1.2e-107;  
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

2 KLVCFYTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLVQEFNG 61  
2 QLTCTYTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLVQEFNG 61  
62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLRYKYSFDCGLDLDWEYP 121  
62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLRYKYSFDCGLDLDWEYP 121  
62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLRYKYSFDCGLDLDWEYP 121  
122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYVDAGYEVDKIAQ 181  
122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYVDAGYEVDKIAQ 181  
122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYVDAGYEVDKIAQ 181  
182 NLDVNLMAVDYDHFSGMEKVTGHSNPLYKQREBSGAASLNVDAAVQOVLKQGTTPASKLI 241  
182 NLDVNLMAVDYDHFSGMEKVTGHSNPLYKQREBSGAASLNVDAAVQOVLKQGTTPASKLI 241  
182 NLDVNLMAVDYDHFSGMEKVTGHSNPLYKQREBSGAASLNVDAAVQOVLKQGTTPASKLI 241  
182 YLDYTHVMTYDLHGSWEGYTGNSPLKYPTDTSNAYLNVDVYNNYKNDGAPAEKLI 241  
242 GMPTVGRSFTLASSSDTRVGAPATGSGTPGPTKGGMLAYEVCWS--KGATKQRIODQ 299  
242 GMPTVGRSFTLASSSDTRVGAPATGSGTPGPTKGGMLAYEVCWS--KGATKQRIODQ 299  
242 GMPTVGRSFTLASSSDTRVGAPATGSGTPGPTKGGMLAYEVCWS--KGATKQRIODQ 299  
300 KVPYIFRDNQWVGDDVESFKTKVSYLKQKGLGGAMVWALDLDLDDFAGFSCNOGRYPILOT 359  
300 KVPYIFRDNQWVGDDVESFKTKVSYLKQKGLGGAMVWALDLDLDDFAGFSCNOGRYPILOT 359

302 EVPYAYQGNVWVGYDNIKSFDIKAQWLKHNFEGAMVWALDLDLDDFTGTFCNOGKFPLIST 361  
360 LQOELSPLPSPGTPPELEVPRKPGQSEP-----EHGPPSPGQDTFCOGKAD 404  
362 LKAL-----GLOSASCTAPAPQIEPIITAAPSGSGSGSGSGSGSGSGGFCFCAVRAN 414  
405 GLYPNPRSSFSYSCAAGRLFOQSCPTGLVFNNSCKCCTW 444  
415 GLYPVANNRNFHWCNVTYQONCQAGLVFDTSCDCCNW 454

RESULT 3  
US-10-004-219B-1  
; Sequence 1, Application US/10004219B  
; Publication No. US20030087414A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase  
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence  
US-10-004-219B-1

Query Match 55.2%; Score 1323; DB 9; Length 476;  
Best Local Similarity 52.2%; Pred. No. 1.2e-107;  
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

2 KLVCFYTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLVQEFNG 61  
23 QLTCTYTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLVQEFNG 82  
62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLRYKYSFDCGLDLDWEYP 121  
62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLRYKYSFDCGLDLDWEYP 121  
62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLRYKYSFDCGLDLDWEYP 121  
122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYVDAGYEVDKIAQ 181  
122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYVDAGYEVDKIAQ 181  
122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYVDAGYEVDKIAQ 181  
182 NLDVNLMAVDYDHFSGMEKVTGHSNPLYKQREBSGAASLNVDAAVQOVLKQGTTPASKLI 241  
182 NLDVNLMAVDYDHFSGMEKVTGHSNPLYKQREBSGAASLNVDAAVQOVLKQGTTPASKLI 241  
182 NLDVNLMAVDYDHFSGMEKVTGHSNPLYKQREBSGAASLNVDAAVQOVLKQGTTPASKLI 241  
203 YLDYTHVMTYDLHGSWEGYTGNSPLKYPTDTSNAYLNVDVYNNYKNDGAPAEKLI 262  
242 GMPTVGRSFTLASSSDTRVGAPATGSGTPGPTKGGMLAYEVCWS--KGATKQRIODQ 299  
242 GMPTVGRSFTLASSSDTRVGAPATGSGTPGPTKGGMLAYEVCWS--KGATKQRIODQ 299  
242 GMPTVGRSFTLASSSDTRVGAPATGSGTPGPTKGGMLAYEVCWS--KGATKQRIODQ 299  
300 KVPYIFRDNQWVGDDVESFKTKVSYLKQKGLGGAMVWALDLDLDDFAGFSCNOGRYPILOT 359  
300 KVPYIFRDNQWVGDDVESFKTKVSYLKQKGLGGAMVWALDLDLDDFAGFSCNOGRYPILOT 359  
323 EVPYAYQGNVWVGYDNIKSFDIKAQWLKHNFEGAMVWALDLDLDDFTGTFCNOGKFPLIST 382  
360 LQOELSPLPSPGTPPELEVPRKPGQSEP-----EHGPPSPGQDTFCOGKAD 404  
383 LKAL-----GLOSASCTAPAPQIEPIITAAPSGSGSGSGSGSGSGSGGFCFCAVRAN 435  
405 GLYPNPRSSFSYSCAAGRLFOQSCPTGLVFNNSCKCCTW 444  
436 GLYPVANNRNFHWCNVTYQONCQAGLVFDTSCDCCNW 475



TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
FILE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
FILE REFERENCE: MEI-030  
CURRENT APPLICATION NUMBER: US/10/097,340  
CURRENT FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: 60/276,025  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/325,149  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/276,026  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/324,967  
PRIOR FILING DATE: 2001/09/26  
PRIOR APPLICATION NUMBER: 60/311,732  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: 60/325,102  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/323,580  
PRIOR FILING DATE: 2001-09-19  
NUMBER OF SEQ ID NOS: 363  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 45

LENGTH: 383

TYPE: PRT

ORGANISM: Homo sapiens

US-10-097-340-45

Query Match 43.4%; Score 1041.5; DB 9; Length 383;  
Best Local Similarity 53.3%; Pred. No. 4.5e-83;  
Matches 195; Conservative 66; Mismatches 96; Indels 9; Gaps 5;

QY 2 KLVCFYTNWAQYRQGEARELPKDLPSLCTHLYAFAGTNHOLSTTEWNETLYOEFG 61  
DB 23 KLVCFYTSWSQYREGDSCFPDADFLCTHLYSPANISNDHIDTWENDVLYGMLNT 82  
QY 62 LKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSDFDGLDWEYP 121  
DB 83 LKKNPNLKTLLSVGGWNFGSKFSKIASNTQSRRTFIKSVPPFLRTHGFDGLDLAWLYP 142  
QY 122 GSQGSFPAVDKERFTTLVODLANAFOEQAOTSGKERLLLSAAVPAQYVDAGYVDKIAQ 181  
DB 143 GR-----DKQHFTTLIKEMKAEFKEAQ-PGKQLLSAALSAGKVTIDSSYDIKISQ 196  
QY 182 NLDVNLMAVDFHGSWEKVTGHNSPLYKQESGAAASLNDAVQOVLQKGTGPASKLIL 241  
DB 197 HLDFTSLTYDFHGAWRQTTHHSPFLRGQGDASDRFSNADYAVSVYVLRGAPANKLVM 256  
QY 242 GMPTGRSFTLASSSDTRVGAPATSGTPPGFTKEGMLAYYEVCSW-KGATKQRIQDQK 300  
DB 257 GIPTGCRSFTLA-SSETGVGAPISGPGIPGRFTKEAGTLAYYEICDFLRGATVHRTLGQQ 315  
QY 301 VPYIFRDQWVGDFDVESPKTKVSYLKQKLGAMVWALDLDDEAGFSCNQG-RYPLIQT 359  
DB 316 VPYATKGNQWVGVDQESVKSVQYLKQKLGAMVWALDLDDEAGFSCNQGLRPLTNA 375  
QY 360 LRQELS 365  
DB 376 IKDALA 381

RESULT 7

US-09-459-749D-17

Sequence 17, Application US/09459749D

Patent No. US20020136716A1

GENERAL INFORMATION:

APPLICANT: Millie, Albert J. T.

TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration

FILE REFERENCE: 0794.016A

CURRENT APPLICATION NUMBER: US/09/459,749D

CURRENT FILING DATE: 1999-12-10

PRIOR APPLICATION NUMBER: 60/111,856

PRIOR FILING DATE: 1998-12-11

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 17

LENGTH: 383

TYPE: PRT

ORGANISM: Sus scrofa

US-09-459-749D-17

Query Match 42.6%; Score 1021.5; DB 10; Length 383;  
Best Local Similarity 52.2%; Pred. No. 2.6e-81;  
Matches 191; Conservative 67; Mismatches 99; Indels 9; Gaps 5;

QY 2 KLVCFYTNWAQYRQGEARELPKDLPSLCTHLYAFAGTNHOLSTTEWNETLYOEFG 61  
DB 23 KLVCFYTSWSQYREGDSCFPDADFLCTHLYSPANISNDHIDTWENDVLYGMLNT 82  
QY 62 LKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSDFDGLDWEYP 121  
DB 83 LKKNPNLKTLLSVGGWNFGSKFSKIASNTQSRRTFIKSVPPFLRTHGFDGLDLAWLYP 142  
QY 122 GSQGSFPAVDKERFTTLVODLANAFOEQAOTSGKERLLLSAAVPAQYVDAGYVDKIAQ 181  
DB 143 GR-----DKRHFTTLVEMKAEFVREA-LPQTERLLLSGAVSAGKVAIDRGYDIAISQ 196  
QY 182 NLDVNLMAVDFHGSWEKVTGHNSPLYKQESGAAASLNDAVQOVLQKGTGPASKLIL 241  
DB 197 HLDFTSLTYDFHGAWRQTTHHSPFLRGQGDASDRFSNADYAVSVYVLRGAPANKLVM 256  
QY 242 GMPTGRSFTLASSSDTRVGAPATSGTPPGFTKEGMLAYYEVCSW-KGATKQRIQDQK 300  
DB 257 GIPTGCRSFTLASSK-TDVGAPASGPGIPGRFTKEGILAYYEICDFLRGATVHRTLGQQ 315  
QY 301 VPYIFRDQWVGDFDVESPKTKVSYLKQKLGAMVWALDLDDEAGFSCNQG-RYPLIQT 359  
DB 316 VPYATKGNQWVGVDQESVKSVQYLKQKLGAMVWALDLDDEAGFSCNQGLRPLTSA 375  
QY 360 LRQELS 365  
DB 376 IKDOLA 381

RESULT 8

US-10-218-743-41

Sequence 41, Application US/10218743

Publication No. US20030096779A1

GENERAL INFORMATION:

APPLICANT: McCall, Catherine A.

APPLICANT: Hunter, Shirley Wu

APPLICANT: Weber, Eric R.

TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS

TITLE OF INVENTION: AND USES THEREOF

FILE REFERENCE: AL-2-C3

CURRENT APPLICATION NUMBER: US/10/218,743

CURRENT FILING DATE: 2002-08-13

PRIOR APPLICATION NUMBER: US/09/292,225

PRIOR FILING DATE: 1999-04-15

PRIOR APPLICATION NUMBER: 60/098,909

PRIOR FILING DATE: 1998-09-02

PRIOR APPLICATION NUMBER: 60/085,295

PRIOR FILING DATE: 1998-05-13

PRIOR APPLICATION NUMBER: 60/098,565

PRIOR FILING DATE: 1998-04-17

PRIOR APPLICATION NUMBER: 09/062,013

PRIOR FILING DATE: 1998-04-17

NUMBER OF SEQ ID NOS: 49

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 41

LENGTH: 490

TYPE: PRT

ORGANISM: Dermatophagoides farinae

US-10-218-743-41

Query Match 31.8%; Score 762; DB 9; Length 490;  
Best Local Similarity 33.3%; Pred. No. 2.2e-58;





Db 93 -EKGVERFNNRLKNPELTTMISLGGWYEGSEKYSDMAANFTYRQQFTQSVDLFLQYK 151  
QY 111 FGLGLDWEYPSQ-GSPAVDKERFTLLVQDLANAFQQAQTSGRERLLLSAAVPAQTY 169  
Db 152 FGLGLDWEYPSQRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAASPGDK 204  
QY 170 VDAGYEVKDIAQNLDVFNLMAYDFHGSWEKVTGHNSPLYKROESGAAS-LNVDAAVQQ 228  
Db 205 IDRAYDIKELNKLFDMMNMYDYHGGWENFYGHNAPLYKRPDETDHLYTFNVNVTMHY 264  
QY 229 WLOKTPASKILGMPTYGRSFTLLASSDTRVGPATGSGTPGPTKEGGMAYVEVCSW 288  
Db 265 YLNNGATDKLVGMVFPFYGRAWSIEDRSKLGDPKAGMSPGFTSGEGVLSYIELCOL 324  
QY 289 KGATKQRIQ-DQ--KVPYIFRDNOMVGFDDVESFKTKVSYLKQKGLGAMVWALDDPFA 345  
Db 325 FQKEWHIQYDEYNAPYGYNDKIWGVYDDLASISCKLAFLKELGVSVMWWSLENDPDK 384  
QY 346 G-----FSCNOG-----RYPLIOTLROELSPLYPST 373  
Db 385 GHCGKPNLLKNVHNMINGDEKNSPECILGSPSTTTPTTTPTTTPTTPTTPTTPTT 444  
QY 374 PELEVP-KPGQSPSEPHGSPGQDT 397  
Db 445 PSPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 469

RESULT 13  
US-10-218-743-18  
; Sequence 18, Application US/10218743  
; Publication No. US20030096779A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine A.  
; APPLICANT: Hunter, Shirley Wu  
; APPLICANT: Weber, Eric R.  
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
; FILE REFERENCE: AL-2-C3  
; CURRENT APPLICATION NUMBER: US/10/218,743  
; PRIOR FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: US/09/292,225  
; PRIOR FILING DATE: 1999-04-15  
; PRIOR APPLICATION NUMBER: 60/098,909  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR APPLICATION NUMBER: 60/085,295  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/098,565  
; PRIOR FILING DATE: 1998-04-17  
; PRIOR APPLICATION NUMBER: 09/062,013  
; PRIOR FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 555  
; TYPE: PRT  
; ORGANISM: Dermatophagoides farinae  
US-10-218-743-18

Query Match 30.8%; Score 738; DB 9; Length 555;  
Best Local Similarity 34.2%; Pred. No. 3.4e-56;  
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;  
QY 2 KLVCYFTNWAQYRQGEARELPKLDPLSLCTHLIYAFAGMTNHOLST-----TEW 50  
Db 34 RIVCYGVTSVTHKVDY-PYTIEDIPFKCTHLMYGFADKIDYKTIQVDFPDYQDDNHNW 92  
QY 51 NDELYQEFNGLKNPKLKTLLAIGGNFQTOKFTDMVATANNRQTFVNSAIRFLKYS 110  
Db 93 -EKGVERFNNRLKNPELTTMISLGGWYEGSEKYSDMAANFTYRQQFTQSVDLFLQYK 151  
QY 111 FGLGLDWEYPSQ-GSPAVDKERFTLLVQDLANAFQQAQTSGRERLLLSAAVPAQTY 169  
Db 152 FGLGLDWEYPSQRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAASPGDK 204

QY 170 VDAGYEVKDIAQNLDVFNLMAYDFHGSWEKVTGHNSPLYKROESGAAS-LNVDAAVQQ 228  
Db 205 IDRAYDIKELNKLFDMMNMYDYHGGWENFYGHNAPLYKRPDETDHLYTFNVNVTMHY 264  
QY 229 WLOKTPASKILGMPTYGRSFTLLASSDTRVGPATGSGTPGPTKEGGMAYVEVCSW 288  
Db 265 YLNNGATDKLVGMVFPFYGRAWSIEDRSKLGDPKAGMSPGFTSGEGVLSYIELCOL 324  
QY 289 KGATKQRIQ-DQ--KVPYIFRDNOMVGFDDVESFKTKVSYLKQKGLGAMVWALDDPFA 345  
Db 325 FQKEWHIQYDEYNAPYGYNDKIWGVYDDLASISCKLAFLKELGVSVMWWSLENDPDK 384  
QY 346 G-----FSCNOG-----RYPLIOTLROELSPLYPST 373  
Db 385 GHCGKPNLLKNVHNMINGDEKNSPECILGSPSTTTPTTTPTTTPTTPTTPTTPTT 444  
QY 374 PELEVP-KPGQSPSEPHGSPGQDT 397  
Db 445 PSPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 469

RESULT 14  
US-09-923-844B-2  
; Sequence 2, Application US/09923844B  
; Patent No. US20020166143A1  
; GENERAL INFORMATION:  
; APPLICANT: Pioneer Hi-Bred International, Inc.  
; APPLICANT: Bao, Zhongmeng  
; APPLICANT: Lu, Guohua  
; TITLE OF INVENTION: Sclerotinia-inducible Genes and  
; FILE REFERENCE: 35718/234631  
; CURRENT APPLICATION NUMBER: US/09/923,844B  
; CURRENT FILING DATE: 2001-08-07  
; PRIOR APPLICATION NUMBER: US 60/224,603  
; PRIOR FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Helianthus annuus  
US-09-923-844B-2

Query Match 20.3%; Score 487; DB 9; Length 371;  
Best Local Similarity 36.1%; Pred. No. 2e-34;  
Matches 126; Conservative 44; Mismatches 135; Indels 44; Gaps 15;  
QY 6 YFTNWAQYRQGEARELPKLDPLSLCTHLIYAFAGMTN-----HQLSTTEWDETLYQ 57  
Db 30 YWPSWAQ-----DFLPSNIQTAYFTHVYAFSLSPNNVTFQFVHRTASALN----- 77  
QY 58 EFN-GLKKNPKLTYLLAIGGNFQT-OKFTDMVATANNRQTFVNSAIRFLKYSFDGLD 115  
Db 78 SFNTALHGKNPPVKTLFSGGSGAGVKQLFSLASSPGSRAAFIRSTIQVARNYFEDGAD 137  
QY 116 LDWEYPSGSGSPAVDKERFTLLVQDLANAFQQAQTSGRERLLLSAAVPAQTYVDAG-- 173  
Db 138 LDWEYPTQ-----TDMNFGULLDEWRVAVNNEATSTGKPRLLLSAATRHBEVRDNGVA 193  
QY 174 -YEVKDIAQNLDVFNLMAYDFHGSW-EKVTGHNSPLYKROESGAAS-LNVDAAVQQMLQ 231  
Db 194 KYPVASINKNLGDNAMCYDYGHTPTPATGAPALYNP-----GSLSTNSGLQSMIS 247  
QY 232 KGTASKILGMPTYGRSFTLLASSDTRVGPATGSGTPGPTKEGGMAYVEVCSWKA 291  
Db 248 AGIQORQKLVGMPLYGWTWKLKNPSVINGIGAPAGIG-PG---NEGAML-YSEVQOQFNAQ 302  
QY 292 TKORI--ODQKVPYI-FRDNOMVGFDDVESFKTKVSYLKQKGLGAMVW 337  
Db 303 NNARVYDTQTVSYISYSGTWTWIGYDDVNSVQRKVQYAKSLNIGGYFFW 351

RESULT 15  
US-09-748-033-3  
; Sequence 3, Application US/09748033  
; Patent No. US20020069431A1  
; GENERAL INFORMATION:  
; APPLICANT: Broadway, Roxanne M.  
; TITLE OF INVENTION: EFFECT OF ENDOCHITINASE AND CHITOSIDASE AND THEIR  
; TITLE OF INVENTION: ENCODING GENES ON PLANT GROWTH AND DEVELOPMENT  
; FILE REFERENCE: 19603/3091  
; CURRENT APPLICATION NUMBER: US/09/748,033  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Streptomyces albidoflavus  
US-09-748-033-3  
  
Query Match : 14.0%; Score 335.5; DB 10; Length 376;  
Best Local Similarity 29.0%; Pred. No. 4e-21;  
Matches 116; Conservative 43; Mismatches 134; Indels 107; Gaps 19;  
  
QY 6 YFTNWAQY-RQGEARFLPKDLPDLCTHLIYAFAGTNNHQLSTTE-----49  
DB 14 YFTWGVGNVHVKNLVTSGSAEKITHINYSFGVGGKCTIGDSFAAYDKAYTAESV 73  
  
QY 50 -----WNDETLYQEFNGLKKNM---PKLTLIAIGWNFGTQKFTDMVATANNRQTFVN 100  
DB 74 DGVADTW-DQPLRGNFNQLEKAKYPHIKVLSFGWTW-SGFETDAV---KNPAFAK 128  
  
QY 101 SAIRFLR--KYS--FDGLDLNWEYPSQG-----SPAVDKERFTTLVQ-----DLA 142  
DB 129 SCHDLVEDPRWADVFDGIDLDWEYPNACGLSCDSGPAALK-----NMVQAMRAQFGTDLV 184  
  
QY 143 N-AFQOEAOQTSGRERLLLSAAVPAGQTYVDAGYEVDKIAQNLDVNLMAVDHGSWEKVT 201  
DB 185 TAAITADASSGGK-----LDAAAYAG-----AAQYFDWYVNTYDFPGAWDK-T 227  
  
QY 202 GHNSPLYKROESG-AAASLNVDAAVQOMLQKTPASKLILGMPTYGRSFTLASSSDTRV 260  
DB 228 GPTAPHSALNSYSGIPKADPHSAAIAKLKAKGVPAKLLIGIGFYGRW-----TGV 280  
  
QY 261 GAPATSGTGPPTKEGGMLAYVEVCSWKATQRTODOKV-----PYIFR 306  
DB 281 TODAPGCTATGPAT-----GTAEAGIEDYKVLKNTCPATGTVGGTAYAKC 325  
  
QY 307 DNOWGVFDDVESFKTVSYLKKGLGGMWALDLDLDDFAG 346  
DB 326 GSNWWSYDTPATIKNTWAKOGLGAFWFPSGDTAGG 365

Search completed: June 29, 2003, 21:13:35  
Job time : 19.1154 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: June 29, 2003, 21:03:33 ; Search time 2892.5 Seconds  
(without alignments)  
4477.355 Million cell updates/sec

Title: US-10-004-219b-10  
Perfect score: 2398  
Sequence: 1 AKLVCFYFTNWAQYRGGEARF.....QQSCPTGLVFSNCKCCTWN 445

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DB=us-10-004-219b-10 -DEVT=xlp  
-Q/cgn2\_1/USPTO.spool/US10004219/runat\_24062003\_160228\_8159/app\_query.fasta\_1.1877  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4 -cdi -LINK=45  
-DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -SCORE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CGN 1.1 6828 @runat\_24062003\_160228\_8159 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_vl.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*  
30: em\_hgt\_hum.\*  
31: em\_hgt\_inv.\*  
32: em\_hgt\_other.\*  
33: em\_hgt\_mus.\*  
34: em\_hgt\_pln.\*  
35: em\_hgt\_rod.\*  
36: em\_hgt\_mam.\*  
37: em\_hgt\_vrt.\*  
38: em\_sy.\*  
39: em\_hgt\_hum.\*  
40: em\_hgt\_mus.\*  
41: em\_hgt\_other.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2398	100.0	1633	9	HSU29615	U29615 Human chito
2	2398	100.0	1636	6	AR206041	AR206041 Sequence
3	2398	100.0	1636	6	AR212138	AR212138 Sequence
4	2398	100.0	1636	6	AX108750	AX108750 Sequence
5	2394	99.8	1643	6	AR172107	AR172107 Sequence
6	2392	99.7	1656	6	AR206042	AR206042 Sequence
7	2392	99.7	1656	6	AR212139	AR212139 Sequence
8	2392	99.7	1656	6	AX108752	AX108752 Sequence
9	2352.5	98.1	1710	9	HSU62662	U62662 Homo sapien
10	2348.5	97.9	1713	6	AR172108	AR172108 Sequence
11	2270.5	94.7	1599	9	AK055165	AK055165 Homo sapi
12	1323	55.2	1625	9	AF290004	AF290004 Homo sapi
13	1309.5	54.6	1526	10	AF154571	AF154571 Mus muscu
14	1309.5	54.6	1538	10	BC011134	BC011134 Mus muscu
15	1309.5	54.6	1557	10	BC034548	BC034548 Mus muscu
16	1304	54.4	1529	4	AB051629	AB051629 Bos tauru
17	1302.5	54.3	1562	5	BJA345054	BJA345054 Bufo japo
18	1301.5	54.3	1530	10	AF290003	AF290003 Mus muscu
19	1072.5	44.7	1369	6	AX059899	AX059899 Sequence
20	1072	44.7	1354	9	AB025008	AB025008 Homo sapi
21	1056	44.0	1801	9	BC008568	BC008568 Homo sapi
22	1053.5	43.9	1056	10	AF062038	AF062038 Rattus no
23	1051	43.8	1741	9	HUMH3G	M80927 Human glyco
24	1046.5	43.6	1433	6	AR042834	AR042834 Sequence
25	1046.5	43.6	1500	9	HSU58515	U58515 Human chiti
26	1046.5	43.6	1526	6	AR042835	AR042835 Sequence
27	1045.5	43.6	1418	9	HSU49835	U49835 Human YKL-3
28	1045.5	43.6	1434	9	HSU58514	U58514 Human chiti
29	1043.5	43.5	1149	6	E01500	E01500 CDNA encodi
30	1043.5	43.5	1149	6	E01501	E01501 CDNA encodi
31	1043	43.5	1449	9	BC011460	BC011460 Homo sapi
32	1030.5	43.0	1733	4	SSU19300	U19300 Sus scrofa
33	1030.5	43.0	1733	4	SSGP38KD	247803 S.scrofa 38
34	1029.5	42.9	1281	4	AY081150	AY081150 Capra hir
35	1015.5	42.3	3946	3	AY051988	AY051988 Drosophil
36	1007	42.0	1641	10	BC004734	BC004734 Mus muscu
37	1007	42.0	1651	10	BC003780	BC003780 Mus muscu
38	1007	42.0	1664	10	BC005611	BC005611 Mus muscu
39	1000	41.7	1616	10	MMBRP39	X93035 M.musculus
40	986.5	41.1	6439	3	AB074977	AB074977 Haemaphys
41	979	40.8	997	4	AF011373	AF011373 Bos tauru
C 42	975.5	40.7	173298	2	AC006489	AC006489 Drosophil
C 43	975.5	40.7	176521	3	AC097725	AC097725 Drosophil
C 44	975.5	40.7	304633	3	AE003477	AE003477 Drosophil
45	970	40.5	1209	10	AY049765	AY049765 Mus muscu

ALIGNMENTS

RESULT 1

```

HSU29615      1633 bp      mRNA      linear      PRI 03-NOV-1995
LOCUS         Human chitotriosidase precursor mRNA, complete cds.
DEFINITION
ACCESSION    U29615
VERSION      U29615.1 GI:1050957
KEYWORDS
SOURCE       Homo sapiens.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1633)
Boot.R.G., Renkema,G.H., Strijland,A., van Zonneveld,A.J. and
Aerts,J.M.
Cloning of a cDNA encoding chitotriosidase, a human chitinase
produced by macrophages
J. Biol. Chem. 270 (44), 26252-26256 (1995)
96064695
PUBMED       7592832
REFERENCE    2 (bases 1 to 1633)
AUTHORS      Boot,R.G.
DIRECT SUBMISSION
TITLE        Submitted (21-JUN-1995) Rolf G. Boot, Department of Biochemistry,
JOURNAL      AMC, E.C. Slater, Institute, University of Amsterdam, Meibergdreef
15, 1105 AZ, Amsterdam, The Netherlands
LOCATION/Qualifiers
1. .1633
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Chi.1"
/cell_type="macrophage"
/clone_lib="macrophage-1 cDNA library of Rolf G. Boot"
13. .1413
/codon_start=1
/product="chitotriosidase precursor"
/protein_id="AAC50246.1"
/db_xref="GI:1050958"
/translation="MYRSVAMAGFMVLLIPMGSAALVCYFTNWAQYRQGEARFLPK
DIDPSLCTHLIYAFAGMTNHSITWENDETLTYEFGNGLKWNPKLKTLLAIGWNFG
TOKFTDMVATANNQTFVNSAIRFLRYKSFGLDLWVEYFGSGSPAVDKERTLLVQ
DLANAFQEOAQTSQKELLISAAVPAGQYVDAGYEVDKIAONLDFVNLMAVDFHGSW
EKTGHNSFLYKQESGSAASLVNDAVQVQLQKGTSPASKILGLMPTVGRSFTLASS
SDTRVGAPATQSPGPFKEGGLMAYEYVCSWKGATKQRIQDQKVPYIFRDNQWGF
DVSEFKTVSYLQKQGLGAMVWALDDFAGFCNQGRIYPLIQLRQELSLPYLPS
GTPELVKPKGQPSPEHGPSQDQTFQCKADGLYPNPRRSRSFYSCAAGRLFQGS
PTGLVNSCKCCTWN"
13. .75
sig_peptide
mat_peptide
76. .1410
/product="chitotriosidase"
/notes="human chitinase"
polyA_signal
polyA_site
1605. .1610
1633
BASE COUNT 354 a 489 c 443 g 347 t
ORIGIN
Alignment Scores:
Pred. No.: 9,34e-176 Length: 1633
Score: 2398.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-004-219B-10 (1-445) x HSU29615 (1-1633)
QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
DB 76 GCAAAACTGTCTGTCTACTTCCACCACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 135
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 136 CTGCCCAAGACTTGGACCCACCCCTTTGACCCACCTATCTACGCCCTTCGCTGCATG 195

```

```

QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 196 ACCAACCCACCCAGCTGAGCACCACTGAGTGAATGACGAGACTCTCTACAGGAGTTCAAT 255
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 256 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGGAGCTGGAAATTC 315
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsuArgGlnThrPheValAsn 100
DB 316 GGCACCTCAGAAGTTTACAGATATGTTAGCCACGGCCACACCCGTCAGACCTTTGTCAAC 375
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 376 TCGGCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGAGTGGAGTAC 435
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 436 CCAGGAAGCCAGGGAGGCCCTCGCTAGCAAGAGCGCTTCACAAACCCCTGGTACAGGAC 495
QY 141 LeuAlaAsnAlaPheGlnGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB 496 TTGGCCATGCTCTTCAGCAGGAGCCACAGACTCAGGGAAGAACCGCTTCTCTGAGT 555
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 556 GCAGCGGTTCCAGCTGGGCAGACCTATGTGGATGCTGGATACAGAGTGAGCAAAATCGCC 615
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 616 CAGAACTCTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTC 675
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu 220
DB 676 ACGGACATACAGCCCTCTACAGAGGCAAGAGAGAGTGTTGCAGCAGCCAGCCTC 735
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
DB 736 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGAGGCCCTGCCAGCAAGCTGATC 795
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB 796 CTTGGCATGCTTACCTACGAGCGCTCTTCACTACCTGGGCTCTCTCATCAGACACCAAGTG 855
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
DB 856 GGGGCCCCAGCCACAGGCTGCGACTCCAGGCCCTTCCAGAGGAGAGGAGGATGCTG 915
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 916 GCCTACTATGAAGTCTGCTCTCTGGAAGGGGGGCCCAACACAGAGAATCCAGGATCAGAAG 975
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
DB 976 GTGCCCTATCATCTCCGGGACCAACAGTGGTGGGCTTTGATGATGGAGAGCTTCAAA 1035
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1036 ACCAAGGTGAGTATCTGAAGCAGAGGAGCTGGGGGGGCCCATGGTCTGGGCATGGAC 1095
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1096 TTAGATGACTTTGGCGGCTTCTCTCTGCAACAGGGCCGATACCCCTCATCCAGACGCTA 1155
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1156 CGGCAGGAACCTGAGTCTTCCATCTTCCAGTCTTCCAGCACCCCGAGCTTGAAGTTCACAAA 1215
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1216 CCAGGTGAGCCCTCTGAACCTGAGCATGGCCCGCCCTGGACAGACACGCTTCTGCCAG 1275
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420

```

Db 1276 GCGAAAGCTGATGGGCTCTATCCCAATCTCGGAAACGGTCCAGCTTCTACAGCTGTGCA 1335  
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
Db 1336 GCGGGGCGGCTGTTCAGCAAAAGCTGCCCGACAGGCGCTGGTGTTCAGCAACCTCTGCAAA 1395  
Qy 441 CysCysThrTIPAsn 445  
Db 1396 TGCTGCACCTGGAAAT 1410

RESULT 2  
AR206041  
LOCUS AR206041 1636 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 1 from patent US 6372212.  
ACCESSION AR206041  
VERSION AR206041.1 GI:21504525  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 1636)  
TITLE Gray, P.W.  
JOURNAL Chitinase: materials and methods  
FEATURES  
Location/Qualifiers  
1..1636  
/organism="unknown"  
BASE COUNT 361 a 491 c 440 g 344 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9.36e-176 Length: 1636  
Score: 2398.00 Matches: 445  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-004-219b-10 (1-445) x AR206041 (1-1636)

Qy 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
Db 65 GCAAAACTGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 124  
Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 125 CTGCCCAAGACTTGGACCCCGCCCTTTGCACCACCTCATCTACGCTTCGGCTGGCATG 184  
Qy 41 ThrAsnHisGlnLeuSerThrThrGluTyrPheAsnAspGluThrLeuTyrGlnGluPheAsn 60  
Db 185 ACCAACCCAGCTGAGCACCCTGAGTGGAAATGACGAGACTCTCTACAGGAGTTCAAT 244  
Qy 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80  
Db 245 GGCTCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAATTC 304  
Qy 81 GlyThrLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
Db 305 GGCACCTCAGAAAGTTTCAGATATGGTAGCCAGCCGCAACACCGTTCAGACCTTTGTCAAC 364  
Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 365 TCGGCCATCATGTTCTTGGCCAAATATACAGCTTTGACGGGCTTGACCTTGAGTGGAGTAC 424  
Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
Db 425 CCAGGAAGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCAACACCCCTGTGTACAGGAC 484  
Qy 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
Db 485 TTGGCCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGGAAGAACGCCCTTCTTCTGAGT 544

Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
Db 545 GCAGCGGTTCCAGCTGGCGAGACCTATGTGATGCTGATGACGAGGTGACAAAATCGCC 604  
Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
Db 605 CAGAACTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTG 664  
Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220  
Db 665 ACGGACATACAGCCCTCTACAGAGGACAGAGAGAGTGGTGCAGCAGCCAGCTC 724  
Qy 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 240  
Db 725 AACGTGGATGCTGTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATC 784  
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
Db 785 CTTGGCATGCCCTTACCTACGGACGCTCTTCACTGGGCTCTCATCAGACACCAAGAGTG 844  
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlyGlyMetLeu 280  
Db 845 GGGGCCCCAGCCAGGCTTGGCACTCCAGGCCCTTCCACCAAGGAGGAGGATGCTG 904  
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 905 GCCTACTATGAAGTCTGCTCTGGAAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 964  
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320  
Db 965 GTGCCCTACATCTTCCGGGACCAACAGTGGGTGGGCTTTGATGATGGAGAGCTTCAAA 1024  
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
Db 1025 ACCAAGCTCAGCTATCTGACGACGAGGAGCTGGGGGGGCGCATGGTCTGGGCACCTGGAC 1084  
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
Db 1085 TTAGATGACTTTTGGCGGCTTCTCTCTGCAACAGGGGCGCATACCCCTCATCCAGACGCTA 1144  
Qy 361 ArgGlnGlnLeuSerLeuProTyrLeuProSerGlyThrProGluLeuValProLys 380  
Db 1145 CGCGAGGAACCTGAGTCTTCCATCTTGGCTTTCAGGACCCCGGAGCTTGAAGTTCCAAA 1204  
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
Db 1205 CCAGGTGACCCCTCTGACCTGAGCATGGCCCGCCCGCCCTGGACACACGTTCTGCCAG 1264  
Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
Db 1265 GCGAAAGCTGATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCA 1324  
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
Db 1325 GCGGGGCGGCTGTTCAGCAAAAGCTGCCCGACAGGCGCTGGTGTTCAGCAACCTCTGCAAA 1384  
Qy 441 CysCysThrTIPAsn 445  
Db 1385 TGCTGCACCTGGAAAT 1399

RESULT 3  
AR212138  
LOCUS AR212138 1636 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 1 from patent US 6399571.  
ACCESSION AR212138  
VERSION AR212138.1 GI:21515643  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 1636)  
TITLE Gray, P.W. and Tjoelker, L.W.  
Chitinase chitin-binding fragments

JOURNAL Patent: US 6399571-A 1 04-JUN-2002;

## FEATURES

Location/Qualifiers

1..1636

/organism="unknown"

BASE COUNT 361 a 491 c 440 g 344 t

## ORIGIN

Alignment Scores:  
Pred. No.: 9 36e-176 Length: 1636  
Score: 2398.00 Matches: 445  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-004-219B-10 (1-445) x AR212138 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
Db 65 GCAAACTGGTCTGCTACTTACCAACTGGGCCAGGTACAGAGGGGGAGGCTCGCTTC 124  
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 125 CTGCCAAGGACTTGGACCCAGCCCTTTGCCACCCATCTACCGCTTCGCTGGCATG 184  
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
Db 185 ACCAACCCAGCTGAGCACCACTGAGTGGAAATGACGAGACTCTTACCGAGGATTCAT 244  
QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80  
Db 245 GGCCTGAAGAAGATGAATCCAAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAATTC 304  
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
Db 305 GGCACCTCAGAAGTTACAGATATGGTAGCCAGCCCAACACCGTCAGACCTTTGTCAAC 364  
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuTyrGluTyr 120  
Db 365 TCGGCATCAGGTTCTGCGCAATACAGCTTTGACGGCTTGACCTTGACTGGGAGTAC 424  
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
Db 425 CCAGGAAGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCAACACCTGGTACAGGAC 484  
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
Db 485 TTGGCCAAATGCCCTTCAGCAGGAGGAGCCAGACCTCAGGGAGGAACCGCTTCTCTGAGT 544  
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
Db 545 GCAGCGGTCCAGCTGGCAGACCTATGTGATGCTGGATACAGGTGACAAATCGCC 604  
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
Db 605 CAGAACCTGGATTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAAGGTC 664  
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220  
Db 665 ACGGACATTAACGCCCTCTCAAGAGGCAAGAGAGAGTGGTCAGCAGCCAGCCCTC 724  
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240  
Db 725 AACGTGGATGCTCTGTGCAACAGTGGCTGCAGAGGGGACCCCTCCAGCAGCTGATC 784  
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
Db 785 CTTGGCATGCTTACCTACGAGCCCTCCTTCAACAGGCTCCTCATCAGACACAGAGTG 844  
QY 261 GlyValProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280  
Db 845 GGGGCCCAAGCCACAGGCTGTGGCACTCCAGGGCCCTTCCCAAGGAAGAGGAGTGCTG 904

QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 905 GCCTACTATGAAGTCTGCTCTCTGGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 964  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
Db 965 GTGCCCTACATCTTCGGGACCAACCAAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1024  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
Db 1025 ACCAAGGTGAGCTATCTGAAGCAGAGGAGCTGGGGGGGCCCATGGTCTGGGCACTGGAC 1084  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
Db 1085 TTAGATGACTTTGCCGGCTTCTCTGCAACACAGGGCCGATACCCCTCATCCAGAGCTA 1144  
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
Db 1145 CGGCAGGAACCTGAGTCTTCATACCTTTCAGGCACCCAGAGCTTGAAGTTCACAAA 1204  
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
Db 1205 CCAGGTGAGCCCTCTGAACCTGAGCATGGCCCGCCCTGGACAAGACACGTTCTGCCAG 1264  
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
Db 1265 GGCMAAGCTGATGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCA 1324  
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
Db 1325 GCGGGCGCGCTTCTCCAGCAAGAGCTGCCCGACAGCGCTGGTTCAGCAACCTCTGCAAA 1384  
QY 441 CysCysThrTrpAsn 445  
Db 1385 TGCTGCACCTGGAAT 1399

## RESULT 4

## LOCUS

AX108750

Sequence 1 from Patent WO0123430.

AX108750

VERSION

AX108750.1

GI:13923942

KEYWORDS

human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1636)

Allison, D.S., Dietsch, G.N., Gray, P.W., Shaw, K.D. and Steiner, B.H.

Human chitinase immunoglobulin fusion proteins

Patent: WO 0123430-A 1 05-APR-2001;

ICOS CORPORATION (US)

## FEATURES

Location/Qualifiers

1..1636

/organism="Homo sapiens"

/db\_xref="taxon:9606"

2..1402

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC37767.1"

/db\_xref="GI:13923943"

## CDS

mat\_peptide

65..1399

/product="unnamed"

BASE COUNT 361 a 491 c 440 g 344 t

## ORIGIN

Alignment Scores:  
Pred. No.: 9,36e-176 Length: 1636  
Score: 2398.00 Matches: 445  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-004-219B-10 (1-445) x AX108750 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
DB 65 GCAAACTGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 124  
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
DB 125 CTGCCCAAGGACTTGGACCCAGCCCTTTGCACCCACCTCATCTACGCCCTTCGCTGGCATG 184  
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
DB 185 ACCAACCCAGCTGAGCACCACTGATGAGTATGAGAGACTCTTACACAGAGTTCAAT 244  
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
DB 245 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAATTTC 304  
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
DB 305 GGCACCTCAGAAGTTCACAGATATGTGTAGCCAGCGGCCAACCAACCGTCAGACCTTTGTCAAC 364  
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
DB 365 TCGGCCATCAGGTTCTTCGGCAATACACTTTCAGCCCTTGACCTTGAGTGGAGATC 424  
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
DB 425 CCAGGAAGCCAGGGAGCCCTCCGTAGACAGGAGCGCTTCAACACCTGGTACAGGAC 484  
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
DB 485 TTGGCCAAATGCCCTTCAGCAGGAAGCCAGACCTCAGGGAAGGAAGCGCTTCTTCTGAGT 544  
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
DB 545 GCAGCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATAGCAGGTGACAAAATCGCC 604  
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
DB 605 CAGAACCTGGATTTGTCAACCTTATGGCTACGACTTCCATGGCTTTGGGAGAAGGTC 664  
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220  
DB 665 ACGGGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGCAGCAGCCCTC 724  
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240  
DB 725 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGACCCCTCCACAGAGCTGATC 784  
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
DB 785 CTTGGCATGCTACCTACGGACGCTCTTCACTACACTGGCTCTCATCAGACACAGAGTG 844  
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280  
DB 845 GGGGGCCCCAGCCAGGGCTGGCACTCCAGGCCCTTCAACAGGAAGAGGAGGATGCTG 904  
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
DB 905 GCCTACTATGAAGTCTGCTCCCTGGAAAGGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 964  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320

DB 965 GTGCCCTACATCTTCCGGGCAACCCAGTGGTGGGCTTGTATGATGTGGAGAGCTTCAAA 1024  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340  
DB 1025 ACCAAGGTGAGCTATCTGAAGCAAGAGGACTGGGGGGGGCATGGTCTGGGCATGGAC 1084  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
DB 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACCCAGGGCCGATACCCCTCATCCAGAGCTA 1144  
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
DB 1145 CGCAGGAACTCAGTCTTCCATCTTGCCTTTCAGGACCCAGAGCTTGAAGTTCCAAA 1204  
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
DB 1205 CCAGTCAAGCCCTTGAACCTGAGCATGGCCCGCCCTGGCAAGACACGCTTCTGCAG 1264  
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
DB 1265 GCAAGAGCTGATGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTTGCA 1324  
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
DB 1325 GCGGGCGGCTTTCAGCAAAAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA 1384  
QY 441 CysCysThrTrpAsn 445  
DB 1385 TGCTGCACCTGGAAT 1399  
RESULT 5  
LOCUS AR172107 1643 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 3 from patent US 6303118.  
ACCESSION AR172107  
VERSION AR172107.1 GI:17911598  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 1643)  
Aerts,J.Maria.Francisus.Gerardus.  
Human chitinase, its recombinant production, its use for  
decomposing chitin, its use in therapy or prophylaxis against  
infection diseases  
JOURNAL Patent: US 6303118-A 3 16-OCT-2001;  
FEATURES  
source Location/Qualifiers  
BASE COUNT 364 a 490 c 442 g 347 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.91e-175 Length: 1643  
Score: 2394.00 Matches: 444  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.83% Indels: 0  
DB: 6 Gaps: 0  
US-10-004-219B-10 (1-445) x AR172107 (1-1643)  
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
DB 79 AAACCTGGTCTGCTACTTCAACAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 138  
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
DB 139 CCCAAGGACTTGGACCCAGCCCTTTCAGCCACCTCATCTACGCCCTTCGCTGGCATGACC 198  
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61

199 AACCCACGAGTGGAGCACCACCTGAGTGGAAATGACAGAGACTCTCTACAGGAGTTCAATGGC 258  
62 LeuLysIleValPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441  
259 CTGAAGAAGATGAATCCCAAGCTTGAGAGACCTCTTACCATCGAGGCTGGAAATTCGGC 318  
82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAlaSer 101  
319 ACTCAGAAGTTACAGATATGTTAGCCACCGCCCAACACCGTCAGACCTTTGTCAACTCG 378  
102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121  
379 GCATCAGGTTTCTGGGCAATACAGCTTTGACGGCTTGTACCTTGGAGTACCA 438  
122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
439 GGAGCCAGGGAGCCCTGCGGTAGACAGGAGCGCTTCAACACCTGGTACAGGACTTG 498  
142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
499 GCCAATGCTTCCAGCAGGAGGCCAGACCTCAGGGAAGGAAGCGCTTCTTCTGAGTGCA 558  
162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
559 GCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACGAGGTGGACAAATCGCCAG 618  
182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201  
619 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTACG 678  
202 GlyHisAsnSerProLeuTyrLysArgGlnGlnGlnSerGlyAlaAlaAlaSerLeuAsn 221  
679 GGACATAACAGCCCTCTACAGAGAGGAGAGAGAGTGGTGCAGCAGCCAGCTCAAC 738  
222 ValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysIleLeu 241  
739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATCCT 798  
242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261  
799 GGATGCTTACCTACGAGAGCTCTCTCACTGGCTCTCTCATCAGACACAGAGTGGGG 858  
262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281  
859 GCGCCAGCAGAGGCTGGCACTCCAGGCCCTTCAACAGGAGGAGGAGTCTGGCC 918  
282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301  
919 TACTATGAAGTCTGCTCTGGAAGGGGGCCACCAACAGAGATCCAGGATCAGAAAGTG 978  
302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThr 321  
979 CCTACATCTTCCGGAGAACCAAGTGGTGGGCTTGTATGATGTGAGAGCTTCAAAACC 1038  
322 LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341  
1039 AAGTTCAGCTATCTGAAGCAGAGGAGCTGGGGGGGCCATGCTTGGCACTGGACTTA 1098  
342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProIleGlnThrLeuArg 361  
1099 GATGACTTTGGGGCTTCTCTGCAACCAAGGGGGCGATACCCCTCATCAGAGCTACGG 1158  
362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGlnValProLysPro 381  
1159 CAGGAAGTGGTCTTCACTACTTGGCTTACGGACCCCAAGAGTTGAGTTCCAAACCA 1218  
382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401  
1219 GGTTCAGCTCTGAACCTGAGCATGGCCCGCCAGCCCTGGCAAGACACGTTCTGCCAGGC 1278  
402 LysAlaAspGlyLeuTyrProAlaProArgGluArgSerSerPheTyrSerCysAlaAla 421  
1279 AAAGCTGATGGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCAGCG 1338

QY 422 GlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441  
Db 1339 GGGCGGCTGTTCCAGCAAGCTGCCGACAGGCTGGTGTTCAGCAACTCTCTGCAATGC 1398  
QY 442 CysThrTrpAsn 445  
Db 1399 TGCACCTGGAAT 1410  
RESULT 6  
AR206042 AR206042 1656 bp DNA linear PAT 20-JUN-2002  
LOCUS Sequence 3 from patent US 6372212.  
DEFINITION AR206042  
ACCESSION AR206042  
VERSION AR206042.1 GI:21504526  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1656)  
AUTHORS Gray, P.W.  
TITLE Chitinase materials and methods  
JOURNAL Patent: US 6372212-A 3 16-APR-2002;  
FEATURES Location/Qualifiers  
1..1656  
source /organism="unknown"  
BASE COUNT 365 a 497 c 447 g 347 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,76e-175 Length: 1656  
Score: 2392.00 Matches: 444  
Percent Similarity: 99.78% Conservative: 0  
Best Local Similarity: 99.78% Mismatches: 1  
Query Match: 99.75% Indels: 0  
DB: 6 Gaps: 0  
US-10-004-219B-10 (1-445) x AR206042 (1-1656)  
QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
Db 90 GCAAACTGGTCTGTCTACTTCCAACTGGGCCAGTACAGACGGGGAGGCTCGCTTC 149  
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 150 CTGCCCCAAGGACTTGGACCCCGAGCTTTGCACCCACCTCATCTACGCTTCGCTGCATG 209  
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
Db 210 ACCAACCCACGACTGAGCACCACCTGAGTGAATGACGAGACTCTCTACAGGAGTTCAAT 269  
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
Db 270 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCTTGTAGCCATCGAGGCTGGAAATTC 329  
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
Db 330 AGCACTCAGAGATTCACAGATATGTTAGCCACCGCCCAACACCGTCAGACCTTTGTCAAC 389  
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 390 TCGGCCATCAGGTTTCTCGCAAAATACAGCTTGTACGGCCCTTGACCTTCACTGGGAGTAC 449  
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
Db 450 CCAGGAAGCCAGGGAGGAGCCCTGCGGTAGACAGGAGCGCTTCAACACCTGCTGACAGAC 509  
QY 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160  
Db 510 TTGGCCATGCTTCCAGCAGGAGGCCAGACCTCAGGGAAGGACGCTTCTTCTGAGT 569  
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180

Db 570 GCAGCGTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATACAGGTGGCAAAATCGCC 629  
Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
Db 630 CAGAACTGGATTGTCAACCTTATGGCTTCAGACTTCCATGGCTCTTGGGAGAGGTC 689  
Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220  
Db 690 ACGGCACATAACAGCCCTCTACAGAGGCAAGAGAGGTGGTGCAGCAGCCAGCTC 749  
Qy 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 240  
Db 750 AACGTGGATGCTGTGCAACAGTGGTGCAGAGGGGAGCCCTCCAGCAAGCTGATC 809  
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgVal 260  
Db 810 CTGGCATGCTTACCTACGAGCGCTCTTCACACTGGCTCTCTATCAGACACAGAGTG 869  
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
Db 870 GGGGCCCCAGCCACAGGGTCTGGCACTCCAGGCCCTTCACCAAGGAAGGAGGTGCTG 929  
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 930 GCCTACTATGAGTCTGCTCTCTGGAAGGGGGGCCACCAACAGAGAATCCAGGATCAGA 989  
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320  
Db 990 GTGCCCTACATCTCCGGGACCAACAGTGGGTGGGCTTTGATGATGGAGAGCTTCAAA 1049  
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
Db 1050 ACCAAGGTCAGCTATCTGAAGCAGAGGAGCTGGGGGGGCTATGCTGGGCACTGGAC 1109  
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
Db 1110 TTAGATGATTTTGGCGGCTCTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1169  
Qy 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
Db 1170 CGCAGGAACTGAGTCTTCATCTTGGCTTCCAGGACCCAGAGCTTGAAGTTCCAAA 1229  
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
Db 1230 CCAGGTTCAGCCCTCTGAACCTGAGCATGGCCCCAGCCCTGGACAGACAGCTTCTGCCAG 1289  
Qy 401 GlyLysAlaAspGlyLeuTyrProLysProLysProArgGluArgSerSerPheTyrSerCysAla 420  
Db 1290 GGCAGAGCTGATGGGCTCTATCCCAATCTCCGGGAACGGTCCAGCTTCTACAGCTGTGCA 1349  
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
Db 1350 GCGGGGCGGCTGTTCAGAAAGTGCCTCCGAGCGGCTGGTGTTCAGCAACTCTTCGCAAA 1409  
Qy 441 CysCysThrTrpAsn 445  
Db 1410 TGCTGCACCTGGAT 1424

## RESULT 7

LOCUS AR212139 1656 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 3 from patent US 6399571.  
ACCESSION AR212139  
VERSION AR212139.1 GI:21515644  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1656)  
AUTHORS Gray,P.W. and Tjoelker,L.W.  
TITLE Chitinase chitin-binding fragments  
JOURNAL Patent: US 6399571-A 3 04-JUN-2002;  
FEATURES Location/Qualifiers

source 1. 1656  
BASE COUNT 365 a 497 c 447 g 347 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,76e-175 Length: 1656  
Score: 2392.00 Matches: 444  
Percent Similarity: 99.78% Conservative: 0  
Best Local Similarity: 99.78% Mismatches: 1  
Query Match: 99.75% Indels: 0  
DB: 6 Gaps: 0  
US-10-004-219B-10 (1-445) x AR212139 (1-1656)  
Qy 1 AlalysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
Db 90 GCAAACTGGTCTCTACTTCCAACTGGGCCCGCAGTACAGACAGGGGGGCGCTCGCTTC 149  
Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 150 CTGCCAAGGACTTGGACCCCGAGCTTTGGACCCACCTCATCTACGCCCTTGGTGGCATG 209  
Qy 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
Db 210 ACCAACCAACAGCTGAGCACCTGAGTGGGAATGACGAGACTCTCTACAGAGTTCAAT 269  
Qy 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
Db 270 GGCCTGAAGAAGATGAAATCCCAAGCTGAAGACCTCTTTAGCCATCGGAGGCTGGAAATTC 329  
Qy 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
Db 330 AGCACTCAAGAAGTTCACAGATATGGTAGCCAGCGCCCAACACCGCTCAGACCTTTGTCAAC 389  
Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 390 TCGGCATCAGGTTTCTCGGCAATACAGCTTTCAGCGCTTGACCTTGACCTGGGAGTAC 449  
Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
Db 450 CCAGGAAGCCAGGGGAGCCCTGCGGTAGACAAGGAGCGCTTCACAAACCTTGGTACAGGAC 509  
Qy 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
Db 510 TTGCCAATGCTTCCAGCAGAGAGCCCGAGCTTCAGGAAGGAACGCTTCTTCTTGAGT 569  
Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
Db 570 GCAGCGGTTCCAGCTGGGCGACACCTATGTGGATGCTGGATACGAGGTGGACAAAATCGCC 629  
Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
Db 630 CAGAACTGGATTTTGTCAACCTTATGGCCTTACGACTTCCATGGCTCTTGGGAGAGGTC 689  
Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220  
Db 690 ACGGCACATAACAGCCCTCTACAGAGGCAAGAGAGTGGTGCAGCAGCCAGCTC 749  
Qy 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 240  
Db 750 AACGTGGATGCTGTGCAACAGTGGCTGCAGAGGGGAGCCCTCCAGCAAGCTGATC 809  
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgVal 260  
Db 810 CTGGCATGCTTACCTACGAGCGCTCTTCACACTGGCTCTCTATCAGACACAGAGTG 869  
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
Db 870 GGGGCCCCAGCCACAGGGTCTGGCACTCCAGGCCCTTCACCAAGGAAGGAGGTGCTG 929  
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300

```
Db      930 GCCTACTATGAAGTCTGCTCTCTGGAGGGGGCCACCAACAGAGAAATCCAGGATCAGAG 989
Qy      301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
Db      990 GTGCCCTACATCTTCGGGACACACAGTGGTGGGCTTTTATGATGTGAGAGCTTCAAA 1049
Qy      321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
Db      1050 ACCAAGTCTAGCTATCTGAAGCAGAGAGGACTGGCGGGGGCCATGCTTGGGCACCTGGAC 1109
Qy      341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuLeuGlnThrLeu 360
Db      1110 TTAGATGACTTTTCCGGCTCTCTCTGCAACAGAGGGCGATACCCCTCATCCAGACGCTA 1169
Qy      361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db      1170 CGCAGGAAGTCTGAGTCTTCTCATACTTTGCTTTCAGGACCCAGAGCTTGAAGTTCCAAA 1229
Qy      381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db      1230 CCAGGTCTAGCTCTGACCTGAGCATGGCCCGCCAGCCCTGGCAAGACAGCTTCTGCCAG 1289
Qy      401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
Db      1290 GGCAGAGCTGATGGCTCTATCCCAATCTCTCGGGAACGGTCCAGCTTCTACAGCTGTGA 1349
Qy      421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db      1350 GCGGGGGGCTGTTCAGCAAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCCAAA 1409
Qy      441 CysCysThrTrpAsn 445
Db      1410 TGCTGCACCTTGAAT 1424
```

```
RESULT 8
LOCUS      AX108752
DEFINITION Sequence 3 from Patent WO0123430.
ACCESSION  AX108752
VERSION     AX108752.1 GI:13923944
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1656)
AUTHORS     Allison,D.S., Dicesch,G.N., Gray,P.W., Shaw,K.D. and Steiner,B.H.
TITLE       Human chitinase immunoglobulin fusion proteins
JOURNAL     Patent: WO 0123430-A 3 05-APR-2001;
            ICOS CORPORATION (US)
FEATURES
     source          1..1656
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             27..1427
                     /notes="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37768.1"
                     /db_xref="GI:13923945"
     translation="MVRSAWAGFWVLLMPGSAKLVICYFTNNAQYQGEARFLPK
     DLDPSLCTHLIYAFAGMTNHLSTTWNDETLQEFNGLKNPKLKLTLAIGWNFS
     TQKFTDMVATANNROTFFVNSAIRFLKYSFGDLDDWEYQSGSPVDKERTFLVQ
     DLANAFQQAQTSQKRLLSAIVPAGQTVDAQYVDKIAQNLDFVNLMAVDFHGSW
     EKVTHNSPLYKROESSGAASLVNDAAVQWLQKGPASKLILGMPTVGRSFTLASS
     SDTRVCAPATGSPFPFFKEGMLAYIEVCWSKGTKORIQDKVPYIFRDNQWVP
     DVSEFKVSYLUKGLGGMWALDDDFAGFCNQGRIPLIQLROELSPLYPIS
     GTEPELVKPGQSPSEHGPSPQDTFCQKADGLYPNPRRSFSYSCAAGRLFOQSC
     PTGLVFSNCKCTWN"
     mat_peptide     90..1424
                     /product="unnamed"
     BASE COUNT      365 a 497 c 447 g 347 t
     ORIGIN
```

```
Alignment Scores:
Pred. No.:      2,76e-175      Length:      1656
Score:          2392.00      Matches:      444
Percent Similarity: 99.78%      Conservative: 0
Best Local Similarity: 99.78%      Mismatches: 1
Query Match:      99.75%      Indels:      0
DB:              6      Gaps:      0

US-10-004-219B-10 (1-445) x AX108752 (1-1656)

Qy      1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
Db      90 GCAAAACTGGTCTGTACTCTTCCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 149
Qy      21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40
Db      150 CTGCCCAAGGACTTGGACCCCGAGCCCTTTGACCCCACTCATCTACGCCCTTCGCTGGCATG 209
Qy      41 ThrAsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
Db      210 ACCAACCAACAGCTGAGCACCACTGAGTGGAAATGACGAGACTCTCTACGAGGAGTTCAAT 269
Qy      61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaLleGlyGlyTrpAsnPhe 80
Db      270 GGCCTGAAAGAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAAATTC 329
Qy      81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db      330 AGCATCTCAGAGATTCACAGATATGTTAGCCAGCGGCCAACACCGTCAGACCTTTGTCAAC 389
Qy      101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db      390 TCGGCCATCAGGTTCTCGCAAAATACAGCTTTCAGCGCTTGACCTTGACCTGGGAGTAC 449
Qy      121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
Db      450 CCAGGAAGCCAGGGAGGCGCTGCGGTAGCAAGAGGCGCTTCACAAACCTGTGTACAGGAC 509
Qy      141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160
Db      510 TTGGCCATGCTTCCAGCAGGAAGCCAGACCTCAGGGAGGAACGCTTCTTCTGAGT 569
Qy      161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
Db      570 GCAGCGGTTCAGCTGGGCGAGACCTATGTGGATGTGGATACAGGTGGACAAAATCGCC 629
Qy      181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
Db      630 CAGAACCTGGATTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTC 689
Qy      201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyValAlaAlaSerLeu 220
Db      690 ACGGACATACAGCCCTCTACAGAGGCGCAGAGAGAGAGTGGTGCAGACCCAGCCTC 749
Qy      221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
Db      750 AACGTGGATGCTGTGTGCAACAGTGGCTGCAGAGGGGAGCCCTCCAGCAGCAAGCTGATC 809
Qy      241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db      810 CTGTGCATCGCTTACCTACGAGCGCTCTTCACACTGGGCTCTCTCATCAGACACCAAGAGTG 869
Qy      261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
Db      870 GGGGGCCCGCAGCACAGGCTTGGCACTCCAGGCCCTTCCACCAAGAGGAGGAGGATGCTG 929
Qy      281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db      930 GCCTACTATGAAGTCTGCTCTCTGGAAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 989
Qy      301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
Db      990 GTGCCCTACATCTTCGGGACACACAGGAGGCTTTGATGATGTGAGAGCTTCAAA 1049
```



QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrrAlaLeuAsp 340  
 DB 1050 ACCAAGTCAGCTATCTGAGCAGAGGAGCTGGCGGGCCATGCTGGGCATGGAC 1109  
 QY 341 LeuAspAspPheAlaGlyPheSerCysLysGlnGlyArgTyrProLeuLeuGlnThrLeu 360  
 DB 1110 TTAGATGATCTTTGGCGGCTCTCTCTGCAACACGAGGCGCATACCCCTCATCCAGACGCTA 1169  
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
 DB 1170 CCGCAGGAACCTGAGTCTTCATCTTTCCTCAGGACCCCGAGAGCTTGAATTCACAAA 1229  
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
 DB 1230 CCAGGTCAGCCCTGAACTGAGCATGCGCCCGCCCTGGACAGACAGACGTTCTGCCAG 1289  
 QY 401 GlyLysAlaAspGlyLeuTyrProLysProLysProArgGluArgSerSerPheTyrSerCysAla 420  
 DB 1290 GGCAGAGCTGATGGGCTCTATCCCAATCTCGGGAACGCTCCAGCTTCTACAGCTGTGCA 1349  
 QY 421 AlaGlyArgLeuPheGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
 DB 1350 GCGGGCGGCTGCTTCCAGAAAGCTGCGCGACAGCGCTGTGTTCAGCAACTCTTGCAAA 1409  
 QY 441 CysCysThrTrrAsn 445  
 DB 1410 TGCTGACCTGGAT 1424

RESULT 9  
 HSU62662 1710 bp mRNA linear PRI 07-SEP-2000  
 LOCUS Homo sapiens chitotriosidase precursor, mRNA, complete cds,  
 DEFINITION alternatively spliced.  
 ACCESSION U62662  
 VERSION U62662.1 GI:9988475  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1710)  
 AUTHORS Boot,R.G., Renkema,G.H., Strijland,A., van Zonneveld,A.J. and Aerts,J.M.  
 TITLE Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages  
 J. Biol. Chem. 270 (44), 26252-26256 (1995)  
 MEDLINE 96064695  
 PUBMED 7592832  
 REFERENCE 2 (bases 1 to 1710)  
 AUTHORS Boot,R.G., Renkema,G.H., Strijland,A. and Aerts,J.M.F.G.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-JUL-1996) Department of Biochemistry, AMC, University of Amsterdam, Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands  
 REFERENCE 3 (bases 1 to 1710)  
 AUTHORS Boot,R.G., Renkema,G.H., Verhoeck,M., Strijland,A., Bliet,J., de Meulemeester,T.M., Mannens,M.M. and Aerts,J.M.  
 TITLE The human chitotriosidase gene. Nature of inherited enzyme deficiency  
 J. Biol. Chem. 273 (40), 25680-25685 (1998)  
 MEDLINE 98421482  
 PUBMED 9748235  
 FEATURES  
 Location/Qualifiers  
 1..1710  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Ch3.1"  
 /cell\_type="macrophage"  
 /clone\_lib="macrophage-1 cDNA library of Rolf G. Boot"  
 10..1173  
 /note="splice variant 40 kDa"  
 /codon\_start=1  
 /product="chitotriosidase precursor"

CDS

/protein\_id="AAG10644.1"  
 /db\_xref="GI:9988476"  
 /translation="MVRSVAMGFMVLLMIWMSAAKLVCVFTWAOVQGEARFLPK  
 DLDPSLTHLIYAFAGMTHSLEWDELTYQEFNGLKMNPKLKLALLAGWNFG  
 TOKFTDMVATANNKQITFVNSAIRFLRKYSFDGLDWEYDPSQSPADVKKERFTLVQ  
 DLNAPQEAQTSKGRLLLSAAVPAQGYVDAGYVDKIAONLDFVNLMDYDFHGSW  
 SKVTGNSPLYKROEESGAASLNVAQOQMLQKGTASKLIILGMPYTGSRFTLASS  
 EDTVGGAPATGSGTGPPTKEGMLAYVEVCWKGATKORIQDQKVPYIFRDNOWVG  
 DVVESFTKVSYLKQKGLGGAMVWALDLDPPGFCNQGRYPLIOTLQELNG"  
 10..72  
 sig\_peptide  
 mat\_peptide 73..1170  
 misc\_feature  
 /product="chitotriosidase"  
 /note="alternatively spliced exon, absent in 50 kDa  
 chitotriosidase (GenBank Accession Number U29615),  
 fragment contains stop codon, therefore coding for a  
 protein of 40 kDa"  
 polyA\_signal 1673..1678  
 polyA\_site 1698..1710  
 BASE COUNT 379 a 503 c 464 g 364 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.19e-172 Length: 1710  
 Score: 2352.50 Matches: 444  
 Percent Similarity: 94.67% Conservative: 0  
 Best Local Similarity: 94.67% Mismatches: 1  
 Query Match: 98.10% Indels: 25  
 DB: 9 Gaps: 1  
 US-10-004-219B-10 (1-445) x HSU62662 (1-1710)  
 QY 1 AlaLysLeuValCysTyrPheThrAsnTrrAlaGlnTyrArgGlnGlyAlaArgPhe 20  
 DB 73 GCAAACTGGTCTGCTACTTCCACAACTGGGCCCGAGTACAGAGGGAGGCTCGCTTC 132  
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
 DB 133 CTGCCCAAGGACTTGGACCCCGAGCCCTTTGCACCCACCTCATCTACGCTTCGCTGGCATG 192  
 QY 41 ThrAsnHisGlnLeuSerThrThrGluTrrAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 DB 193 ACCAACCAACAGCTGAGCACCACCTGAGTGGATACGAGACTCTCTACAGGAGTTCAT 252  
 QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrrAsnPhe 80  
 DB 253 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAAATTC 312  
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
 DB 313 GGCACCTCAGAAGTTCCACAGATATGCTAGCCAGGCCAACCAACCGTCAGACCTTTCTCAAC 372  
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrrGluTyr 120  
 DB 373 TCGGCCATCAGGTTTCTCGCAAAATACAGCTTGTGACGCCCTTGACCTTGACCTGGAGTAC 432  
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
 DB 433 CCAGNAGCCAGGGAGGCCCTGCGGTAGCAAGGAGCCCTTCACAAACCTCGTGTGACAGAC 492  
 QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
 DB 493 TTGGCCATGCTTCCAGCAGGAGGCCAGACCTCAGGGAAGGAACGCTTCTCTGAGT 552  
 QY 161 AlaAlaValProAlaGlyGlnThrTrrValAspAlaGlyTyrGluValAspLysIleAla 180  
 DB 553 GCAGCGGTTCAGCTGGGCGAGACCTATGTGGATGTGTGATACGAGGTGACAAAATCGCC 612  
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTrrAspPheHisGlySerTrrGluLysVal 200  
 DB 613 CAGAACCTTGGATTTTGTCAACCTTATGGCCTACGACTTCCATGGCTCTTGGGAGAGGTC 672  
 QY 201 ThrGlyHisAsnSerProLeuTrrLysArgGlnGluGluSerGlyAlaAlaAlaSerLeu 220

Db 673 ACGGACATAACAGCCCTCTACAAGAGCGCAAGAGAGAGTGGTGCAGCAGCCAGCCTC 732  
Qy 221 AsnValAspAlaalaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 240  
Db 733 AACGTGGATGCTGCTGTGAACAGTGGCTGCAAGAGGGACCCCTCCAGCAGCTGATC 792  
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
Db 793 CTGGGATGCTTACCTACGAGCGCTCTTCAACAGTGGCTCTCTATCATCAGACCCAGAGTG 852  
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280  
Db 853 GGGGCCCCAGCCAGCGCTGCGCATCTCCAGGGCCCTTCAACAGAGAGAGGGATGCTG 912  
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgileGlnAspGlnLys 300  
Db 913 GCGTACTATGAAGTCTGCTCTGGAAGGGGGGCCACCAACAGAGAAATCCAGGATCAGAAG 972  
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
Db 973 GTGCCCTACATCTTCGGGACACACAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1032  
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
Db 1033 ACCAAGGTCACTATCTGAAGCAGAGAGGAGTGGGGGGGCCATGCTCTGGCAGCTGGAC 1092  
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuLeuGlnThrLeu 360  
Db 1093 TTAGATGACTTTGGCCGGCTTCTCCTGCACACAGGGGCCGATACCCCTCATCCAGAGCTA 1152  
Qy 361 ArgGlnGluLeuSer----- 365  
Db 1153 CGCGAGGAAGCTGAA-TGGGTAAAGCCTTAACCTGCTGCATGTGAGGCCAGGTGTTC 1211  
Qy 366 -----LeuProTyrLeuProSerGlyThrProGluLeu 376  
Db 1212 CTGTGGCACTGTCTTTCAGCTGTAGGCTCTTCCATATCTTCCAGCCACCCAGAGCTT 1271  
Qy 377 GluValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAsp 396  
Db 1272 GAAGTTCCAAACACAGTCAAGCTCTGAACTGAGCATGGCCCCAGCCCTGGACAAAGAC 1331  
Qy 397 ThrPheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerPhe 416  
Db 1332 ACCTTCTGCAGGCAAGCTGATGGCTCTATCCCAATCTCTCGGAAACGGTCCAGCTTC 1391  
Qy 417 TyrSerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSer 436  
Db 1392 TACAGCTGTGCAGCGGGGGCTGTTCAGCAAGCTGCCCAAGCTGCCAGAGCCCTGGTTCAGC 1451  
Qy 437 AsnSerCysLysCysCysThrTrpAsn 445  
Db 1452 AACTCTGCAATGCTGCACCTGGAAT 1478

RESULT 10  
LOCUS AR172108  
DEFINITION Sequence 5 from patent US 6303118.  
ACCESSION AR172108  
VERSION AR172108.1 GI:17911599  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1713)  
AUTHORS Aerts,J.Maria.Francisus.Gerardus.  
TITLE Human chitinase, its recombinant production, its use for  
decomposing chitin, its use in therapy or prophylaxis against  
infection diseases  
JOURNAL Patent: US 6303118-A 5 16-OCT-2001;  
FEATURES Location/Qualifiers  
source 1..1713

BASE COUNT 380 a 504 c 464 g 365 t  
ORIGIN  
Alignment Scores:  
Score: 6.49e-172 Length: 1713  
Pred. No.: 2348.50 Matches: 443  
Percent Similarity: 94.66% Conservative: 0  
Best Local Similarity: 94.66% Mismatches: 1  
Query Match: 97.94% Indels: 25  
DB: 6 Gaps: 1  
US-10-004-219B-10 (1-445) x AR172108 (1-1713)  
Qy 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
Db 79 AAACGTGGTCTGCTACTTCCACCACTGGGCCCCAGTACACAGAGGGGAGGCTCGCTTCCTG 138  
Qy 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
Db 139 CCCAAGGACTTGGACCCCGAGCTTTGCACCCACCTCATCTACGCCCTTCGCTGGCATGACC 198  
Qy 42 AsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
Db 199 AACCCACAGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTCTACAGAGATTCAATGGC 258  
Qy 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81  
Db 259 CTGAAGAGATGAATCCCAAGCTGNAGACCTGTGTAGCCATCGGAGGCTGGAAATTCGGC 318  
Qy 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
Db 319 ACTCAGAAGTTTCACAGATATGCTAGCCACGCGCAACACCGTCAGACCTTTGTCAACTCG 378  
Qy 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTrpPro 121  
Db 379 GCCATCAGGTTTCTGCGCAATATACAGCTTTCAGCGCTTGACCTTGACTGGGAGTACCCA 438  
Qy 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
Db 439 GGAAGCCAGGGGAGCCCTGCTCCCTAGACAAGAGCGCTTCAACACCTCGTGCAGGACTTG 498  
Qy 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
Db 499 GCGAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGGAACGCTTCTTCTGTAGTGCA 558  
Qy 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
Db 559 GCGGTTCCAGCTGGGAGACCTATGTGGATGACGAGGTGGACAAATATGCCAG 618  
Qy 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201  
Db 619 AACCTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGAGAGGTACAG 678  
Qy 202 GlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyAlaAlaSerLeuAsn 221  
Db 679 GGACATAACAGCCCTCTACAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCCCTCAAC 738  
Qy 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241  
Db 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAGCTGATCTT 798  
Qy 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261  
Db 799 GGCATGCCCTACCTAGGAGCGCTCTTCACTGGCTCTCTCATCAGACACAGAGGTGGG 858  
Qy 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281  
Db 859 GCGCCAGCCACAGGGTCTGGCAGCTCCAGGCCCTTCCACAGGAAGAGGAGGTGCTGCC 918  
Qy 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgileGlnAspGlnLysVal 301  
Db 919 TACTATGAAGTCTGCTCTCTGGAAGGGGGCCCAACAGAGATCCAGATCAGAGGTG 978

HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

## FEATURES

source  
Location/Qualifiers  
1..1599  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CD34C2000051"  
/cell\_type="CD34+ Cells"  
/clone\_lib="CD34C2"  
/note="Cloning vector: pME18SFL3-primary culture, CD34+ Cells"

BASE COUNT 339 a 481 c 437 g 342 t  
ORIGIN

## Alignment Scores:

Pred. No.: 6,12e-166 Length: 1599  
Score: 2270.50 Matches: 425  
Percent Similarity: 95.51% Conservat: 0  
Best Local Similarity: 95.51% Mismatches: 1  
Query Match: 94.68% Indels: 19  
DB: 9 Gaps: 1

US-10-004-219B-10 (1-445) x AK055165 (1-1599)

QY 302 ProTyrllePheArAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321  
Db CCTACATCTTCGGGACACACAGTGGGTGGCTTGATGATGGAGAGCTTCAAAACC 1038

QY 322 LysValSerTyrlleLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341  
Db AAGGTGAGTATCTGAAGCAGAGGAGTGGGGGGGGCCCATGCTGGGCACCTGAGACTTA 1098

QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrlleLeuThrLeuArg 361  
Db GATGACTTTGGCGGCTTCTCTGCAACAGGCGCCGATACCCCTCATCAGACGCTACGG 1158

QY 362 GlnGluLeuSer----- 365  
Db 1159 CAGGAAGTAAAGCTTAACTGCCTGTGCATGTGAGGCCAGGTGTTGCTCTG 1217

QY 366 -----LeuProTyrlleProSerGlyThrProGluLeuGlu 377  
Db 1218 TGGCAGCTGTGCTTACAGTGTAGGTCTTCCATCTTCCCTTTCAGGCACCCAGAGCTTGA 1277

QY 378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397  
Db 1278 GTTCAAAACACAGGTGAGCCCTCTGAACCTGAGCATGGCCCAAGCCCTGGCAAGACAGC 1337

QY 398 PheCysGlnGlyLysAlaAspGlyLeuTyrlleProAsnProArgGluArgSerPheTyrlle 417  
Db 1338 TTCTGCCAGGCAAGCTGATGGCTCTATCCATCTCTGGGACCGTCCAGCTTCTAC 1397

QY 418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437  
Db 1398 AGCTGTGACGGGGGGCGGTGTTCAGCAAAAGCTGCCGACAGGGCTGGTGTTCAGCAAC 1457

QY 438 SerCysLysCysCysThrTrpAsn 445  
Db 1458 TCTGCAATGCTGCACCTGGAAT 1481

## RESULT 11

AK055165 1599 bp mRNA linear PRI 01-AUG-2002  
LOCUS Homo sapiens CDNA FLJ30603 fis, clone CD34C2000051, highly similar  
DEFINITION to Human chitotriosidase precursor mRNA.  
ACCESSION AK055165  
VERSION AK055165.1 GI:16549834  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens CD34+ Cells cDNA to mRNA, clone\_lib:CD34C2  
clone:CD34C2000051.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1  
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,  
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,  
Kaminara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,  
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,  
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,  
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,  
Kikuchi, H., Murakawa, K., Kaneshiro, K., Takahashi-Fujii, A.,  
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,  
Nagahari, K., Masuno, Y., Nagai, K. and Sugiyama, T.  
NEDO human cDNA sequencing project

## TITLE

Unpublished

## JOURNAL

2 (bases 1 to 1599)

## REFERENCE

Isogai, T., Otsuki, T. and Sugiyama, T.  
Direct Submission  
Submitted (24-Oct-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

```
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db 763 CTTGGGATGCTTACCTACGGAGCGCTCTTACACACTGGCTCTCTCATCAGACACAGAGTG 822
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
Db 823 GGGGCCACAGCCACAGGGTGTGGCACTCCAGGCGCCCTTACCAGAGGAGGAGGATGCTG 882
QY 281 AlaTyrTyrGluValCysSerTyrLysGlyAlaThrLysGlnArgLysGlnAspGlnLys 300
Db 883 GCTACTATGAAGTCTGCTCTCGAAGGGGGGGCCACCAACAGAGAGATCCAGGATCAGAAG 942
QY 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspAspValGluSerPheLys 320
Db 943 GTGCCCTACATCTTCCGGGACACACAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1002
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
Db 1003 ACCAAGTCTAGCTATCTGAAGCAGAGAGGAGTGGGGGGGCCCATGTGCTGGGCACCTGGAC 1062
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1063 TTAGATGACTTTGCCGGCTTCTCTCGAACACAGGGCCGATACCCCTCATCCAGAGCTA 1122
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1123 CGGCGAGAACTGAGTCTTCCATACTTGCCTTCAGGCACCCAGAGCTTGAAGTTCCAAA 1182
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1183 CCAGGTTCAGCCCTCTGAACCTGAGCATGGCCCGCCCTGGACAGACACGCTTCTGCCAG 1242
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
Db 1243 GGCAAGCTGATGGGCTCTATCCATCTCCGGAACGGTCCAGCTTCTACAGCTGTGCA 1302
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db 1303 GGGGGCGGCTGTTCAGCAAAAGCTGCCGACAGGCTGGTGTTCAGCAACTCTCTGCAA 1362
QY 441 CysCysThrTrpAsn 445
Db 1363 TGCTGCACCTGGAAT 1377

RESULT 12
AF290004
LOCUS AF290004
DEFINITION Homo sapiens acidic mammalian chitinase precursor, mRNA, complete cds.
ACCESSION AF290004
VERSION AF290004.1 GI:12597292
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1625)
AUTHORS Boot,R.G., Blommaert,E.F., Swart,E., Ghausharali-van der Vlugt,K.,
Bijsl,N., Moe,C., Place,A. and Aerts,J.M.
TITLE Identification of a novel acidic mammalian chitinase distinct from
chitotriosidase
J. Biol. Chem. 276 (9), 6770-6778 (2001)
MEDLINE 21125893
PUBMED 11085997
REFERENCE 2 (bases 1 to 1625)
AUTHORS Boot,R.G., Verhoeck,M., Swart,E. and Aerts,J.M.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical
Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105
AZ, The Netherlands
FEATURES
source Location/Qualifiers
1..1625
/organism="Homo sapiens"
```

## CDS

```
/db_xref="taxon:9606"
/tissue_type="stomach; lung"
104..1534
/EC_number="3.2.1.14"
/Note="chitinase precursor; AMCase; chitinase activity
demonstrated by expression in COS cells; similar to
sequences presented in GenBank Accession Numbers AB025008
and AB025009"
/codon_start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60019.1"
/db_xref="GI:12597293"
/translation="MTKLILLTGLVLLNLQSGAYQLTCYFTNWAQYRPGLRMPD
NIDPLCTHLIYAFAGRNQNEITTEWNTVLYQAFNGLKNKSLQLTLAIGWNFG
TAPFTAMYSTPENQTFITSVIKPLROYEFDGLDFDWEYPGSRGSPPODKHLFTLVLSQ
EMRAFEQEAQKQINKPRIMVAAVAAGISNIQSGVEIPLSQYLDYIHWTVDLHGSW
EYTGENSEPLKYKPYDITGSNAYLVNDVYMNWKONGAPAEKLIYGFPTYGHNFIILSNP
SNTGSGAGGAGPAGPYAKESGWIWAYEIEITFLKNGATQWDAPQEVPIAYQGNWV
GYDNIKSPDIKAQWLKHNKFGMAWVADLDDFTGTCNQKGFPLITLTKALGLQSA
SCTPAQPIETITAAPSGSGNGSGSSSGSGSGSGGFCVAVRANGLYPVANRRNAFMC
VNGVTYQNCQAGLVFDTSDCCNWA"
sig_peptide 104..166
BASE COUNT 387 a 468 c 399 g 371 t
ORIGIN
Alignment Scores:
Pred. No.: 6,75e-93 Length: 1625
Score: 1323.00 Matches: 240
Percent Similarity: 67.3% Conservative: 70
Best Local Similarity: 52.1% Mismatches: 126
Query Match: 55.1% Indels: 24
DB: 9 Gaps: 3
US-10-004-219B-10 (1-445) x AF290004 (1-1625)
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
Db 170 CAGCTGACATGCTACTTACCACCTGGGCCAGTACCGGCCAGGCTGGGGCGCTTCATG 229
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
Db 230 CTGCAACAATCGACCCCTGCTCTGTACCCACCTGATCTACGCTTTGCTGGGAGGAG 289
QY 42 AsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
Db 290 AACACGAGATCACCAACCATCGAATGGAACGATGACTCTCTACCAAGCTTCAATGCG 349
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnGly 81
Db 350 CTGAAAAATAAGAACAGCCAGCTGAAAACTCTCTGGCCATTTGGAGGCTGGAACTTCGGG 409
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
Db 410 ACTGCCCTTTTCACTGCCATGTTTCTACTCTGTAGAACCCGACAGCTTTCATCCTCA 469
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
Db 470 GTCATCAAAATTCCTGGCCAGTATGAGTTTGACGGCTGGAGCTTTGACTGGGAGTACCT 529
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
Db 530 GGCTCTCGTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGTCTGTGTGGAGAAATG 589
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
Db 590 COTGAAGCTTTTGGACAGAGCCCAAGCAGATCAACAAGCCAGGCTGATGCTGCTGCT 649
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
Db 650 GCAGTAGCTGTGTCATCTCCAATATCCAGTCTGCTGATGATGATGATGATGATGATGAT 709
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
```

Db 710 TACCTGGACTACATCATGCTACGACCTACGACCTCCATGGCTCTCTGGAGGCGCTACACT 769

QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsn 221

Db 770 GGAGAGACAGCCCTCTACAAATACCGACTGACACCGGACGACCGCTACCTCAAT 829

QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241

Db 830 GTGGATTATGTCATCACTACTGGAAGGACATGAGCACCAGCTGAGAGCTCATCGTT 889

QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGly 261

Db 890 GGATTCCTTACCTATGGACACAACTTCATCTCGAGCAACCCCTCCACACTGGAATGCT 949

QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlnGlyGlyMetLeuAla 281

Db 950 GCCCCACCTCTGGTGTCTGCTGGCTGAGGCTATGCCAGAGTCTGGGATCTGGGCT 1009

QY 282 TyrTyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgGlnAspGln 299

Db 1010 TACTACGAGATCTGACTCTCTGAAATGAGCAGCAGCTACGGATGGATGCCCTCAG 1069

QY 300 LysValProTyrIlePheArgAspGlnTrpValGlyPheAspAspValGluSerPhe 319

Db 1070 GAAGTGCCTTATGCTCTATCAGGCGCAATGTGGTGGCTATGACAACTCAAGAGCTTC 1129

QY 320 LysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeu 339

Db 1130 GATATTAGGCTCAATGGCTTAAGCAACAATAATTGGAGGCGCCATGCTCTGGGCCATT 1189

QY 340 AspLeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThr 359

Db 1190 GATCTGGATGACTCTCACTGGCACTTCTGCACAGGCGCAAGTTCCTCTATCTCCACC 1249

QY 360 LeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro 379

Db 1250 CTGAAGAGGCGCTC-----GGCTGCAGAGTCAAGTTGCACG 1288

QY 380 LysProGlyGlnProSerGluPro----- 387

Db 1289 GCTCCAGCTCAGCCCATGAGCAATACTGCTGCCAGTGGCAGCGGGAACGGGAGC 1348

QY 388 -----GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAsp 404

Db 1349 GCGAGTAGCAGCTCTGAGGCGAGCTCGGAGCAGTGATCTGCTCTCAGAGCCAC 1408

QY 405 GlyLeuTyrProAsnProArgGluArgSerPheTyrSerCysAlaAlaGlyArgLeu 424

Db 1409 GGCCTCTACCCGCTGGCAATAACAGAAATGCTCTGCGCACTGGCTGAATGGAGTCA 1468

QY 425 PheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp 444

Db 1469 TACCAGCAAACTCGCAGCGCGGCTTGTCTCGACACCAGCTGTGATTGCTGCAACTGG 1528

RESULT 13

AF154571

LOCUS 1526 bp mRNA linear ROD 08-FEB-2000

DEFINITION Mus musculus putative chitinase precursor (YNL) mRNA, partial cds.

ACCESSION AF154571

VERSION AF154571.1 GI:6934189

KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1526)

PRICE, P.A., Harris, S.C. and Williamson, M.K.

YNL, A Putative Mouse Chitinase

Unpublished

2 (bases 1 to 1526)

PRICE, P.A., Harris, S.C. and Williamson, M.K.

Direct Submission

Submitted (26-MAY-1999) Biology, University of California, San

Diego, 9500 Gilman Drive, Mail Code 0368, La Jolla, CA 92093-03687, USA

FEATURES

source

Location/Qualifiers

1..1526

/organism="Mus musculus"

/strain="ATCC147668"

/db\_xref="ATCC:147668"

/db\_xref="taxon:10090"

/clone="IMAGE 1209227"

/tissue\_type="skin"

<1..1526

/gene="YNL"

<1..1419

/gene="YNL"

/codon\_start=1

/product="putative chitinase precursor"

/protein\_id="AAF31644.1"

/db\_xref="GI:6934190"

/translation="AKLLLVTLALLNAQLGSAYNLICYFTNMAQYRPGLSFXPDD INPCLTHLIYAFAGMONEITTEWINDVLYKAFNDLKRNLSKLKTLAIGGMVFGT APETTWSTSONROTFTSVIKFLOYGFDGLDLDWEYPGSRGSPDQKHLPVLVKE MREAPEQAEIESNRPRLMVTAAVAGISNIQAGYEIPELSKYLDPFIHMTVDLHGSWE GITGENSFYKIPITETGSNAYLNVDYVNNYKNGAPAEKLIIVGPEPCHTIFLRNPS DNGIAPSGDGPAGPYTRQAGFWAYEICTFLRSGATEVMDASOEVPYAKANEMWG YDNIGKSFVKAQWLKQNNFGAMTWAIDDDFTGSCFCDQGFPLTSTLNKALGISTEG TADVPSEPVTPPGSGSGSGSGSGSGSGGFCADKADGLYPVADDRNAFWQCINGI TYQHQCOAGLVFDTSNCNWP"

sig\_peptide 41..60

/gene="YNL"

mat\_peptide 61..1416

/gene="YNL"

misc\_feature 394..417

/product="putative chitinase"

/gene="YNL"

/note="putative conserved region; active site"

misc\_feature 415..417

/gene="YNL"

/note="putative invariant Glu; active site"

BASE COUNT 383 a 381 c 364 g 398 t

ORIGIN

Alignment Scores:

Pred. No.: 6,846-92 Length: 1526

Score: 1309.50 Matches: 236

Percent Similarity: 67.32% Conservative: 71

Best Local Similarity: 51.75% Mismatches: 128

Query Match: 54.61% Indels: 21

DB: 10 Gaps: 3

US-10-004-219B-10 (1-445) x AF154571 (1-1526)

QY 3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22

Db 67 CTGATATGCTATTTCCACCACTGGGCCAGTATCGGCCAGCTCTGGGAGCTTCAAGCT 126

QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42

Db 127 GATGACATTAACTCCCTGGCTGTGACTCACCGATCTATGCTTGTCTGGGATGAGAAC 186

QY 43 HisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62

Db 187 AATGAGATCACCACTAGATGAATGATGTTACTCTCTATAAAGCTTCAATGACTTG 246

QY 63 LysLysMetAsnProLysLysThrLeuLeuAlaIleGlyTyrAsnPheGlyThr 82

Db 247 AAAAAACAGAACACAAACCCCTCTGGCAATTTGGAGGCTTGAACCTTTGGAAC 306

QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102

Db 307 GCTCCTTTTACCTACCATGTTTCCACTTCTCAGAACCCGACACCTTCAATACCTCAGTC 366

QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122

Db 367 ATCAAAATTTTCGCTCAGTATGGTTTGTATGGACTGGACCTGGAGTACCGGAAATACCCAGGC 426  
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142  
Db 427 TCAGTGGGAGCCCTCTCAGACAGCATCTCTTCACTGTCTCTGTGAGGAAATGCCT 486  
QY 143 AsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAlaAla 162  
Db 487 GAAGCTTTTTCAGCAGGAGGCTATTGAGAGCAACAGGCCAGCATGATGTTACTGTCTGCT 546  
QY 163 ValProAlaGlnThrThrValAspAlaGlyTyrGluValAspLysLeuAlaGlnAsn 182  
Db 547 GTAGCTGGTGGGATTTCCAACTCCAGCTGGCTATGAGATCCCTCAACTTTCTAGTAC 606  
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTyrGluLysValThrGly 202  
Db 607 CTGGATTTTCATCATGTCATGATATGACCTCATGCTCTCTGGAGGGCTACACTGGG 666  
QY 203 HisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222  
Db 667 GAGATAGTCTCTTTACAAATACCTCTAGACTGGTAGCAATGCCTACCTCAATGTG 726  
QY 223 AspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242  
Db 727 GATTATGTCATGAACATTGGAAGAACAAATGGAGCCCGAGCTGAGAAAGCTCATTTGTA 786  
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262  
Db 787 TTCCAGAGTATGGACACACCTTCATCTCTGAGAACCCCTCTGATATGGAATGTGTGCC 846  
QY 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGlyGlyMetLeuAlaTyr 282  
Db 847 CCTACCTCTGGTATGCGCTGTGGCGCTCTATACCAGACAGCTGGTCTTGGGCGCTAC 906  
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 907 TATGAGATTTGACCTTTCTGAGAGTGGAGCCAGCTGAGGTCTGGGATGCTCCCAAGAA 966  
QY 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspAspValGluSerPheLys 320  
Db 967 GTGCCCTATGCTTAAAGCCACAGAGTGGTGGCTGATGACATATCAAGAGCTTCAGT 1026  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
Db 1027 GTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCAATTGAC 1086  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
Db 1087 CTTGATGACTTCACTGGCTCTTCTGTGATCAGGAAATTTCTCTGACTTCTACTTTG 1146  
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
Db 1147 ACAAGAGCCCTGGCATATCCACTGAAGGTTCACAGCTCTCTGACGTG-----Glu 388  
QY 381 ProGlyGlnProSerGluPro-----CCTTCGAGCAGTACTCTCTCCAGAGTGGGAGTGGGAAGC 1245  
Db 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyAlaAspGlyLeuTyrPro 408  
Db 1246 TCCGGAGGAGCTCTGGAGCAGTGAATCTGTGCGCAGAACAGCAGATGGCTCTACCT 1305  
QY 409 AsnProArgGluArgSerSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428  
Db 1306 GTGGCAGATCAGAGAATGCTTTTGGCAGTGCATCAATGGAATCATACACAGCAGCAT 1365  
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp 444  
Db 1366 TGTCAAGCAGGCGCTGTTTTCATACACAGCTGTAATTTGTCGAACGTG 1413

RESULT 14

BC011134

LOCUS

BC011134 1538 bp mRNA linear ROD 07-AUG-2002

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REMARK

## COMMENT

## FEATURES

## source

## CDS

## BASE COUNT

## ORIGIN

## Alignment Scores:

## Pred. No.:

## Score:

## Percent Similarity:

## Best Local Similarity:

## Query Match:

## DB:

## US-10-004-219B-10 (1-445) x BC011134 (1-1538)

## QY

## Db

Mus musculus, similar to eosinophil chemotactic cytokine, clone  
MGC:18771 IMAGE:4165150, mRNA, complete cds.

BC011134

BC011134.1 GI:15029821

MGC.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1538)

Strausberg, R.

Direct Submission

Submitted (25-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)

Goonaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 24 Row: e Column: 11.

Location/Qualifiers

1..1538

/organism="Mus musculus"

/db\_xref="taxon:10090"

/map="FVB/N"

/clone="MGC:18771 IMAGE:4165150"

/tissue\_type="Salivary gland, 10 week old female mouse"

/clone\_lib="NCI CGAP\_SG2"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

323..1420

/codon\_start=1

/product="Similar to eosinophil chemotactic cytokine"

/protein\_id="AAH11134.1"

/db\_xref="GI:15029822"

/translation="MVSTQRQTFITSVIKPLROYGFDGLDLDWEYFGSRGSPQDK  
HLFTLVKEMREAFQEAEIESNPRLMTAAVAGGISNIOAGYEIPELSKYLDFHVM  
TYDLHSGVETGENSPLYKPTETGSNAYLVNVMYKNGAPAEKLIYGFPEYG  
HTFLRNPSDNGIGAPTSGDGPAGPYTFOAGFWAYEICTFLRSGATVWDASQVRY  
AYKANELGYDNISFSKAKWLQNNFGNMIWALDLDGTSFGDQKPLTSTLN  
KALGISTEGCTAPDVPSEPTTPPGSGGGSSGGSGGSCGFCADKADGLYPVADRN  
AFQCGINTYQQHQAGLVDFDTCNCNWP"

393 a 382 c 365 g 398 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-004-219B-10 (1-445) x BC011134 (1-1538)

QY

Db

3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22

68 CTGATATGCTATTTTCCACCACTGGGCCCGCCAGTATCGCCAGGCTCTGGGAGCTTCAAGCCT 127

QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42  
DB 128 GATGACATTAAACCCCTGCTGTACTACCTGATCTATGCTCTTTGGGATGAGAAC 187  
QY 43 HisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62  
DB 188 AATGAGATCACCAACCATAGAATGGAATGATGATCTCTCTATAAAGCTTTCAATGACTTG 247  
QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr 82  
DB 248 AAAACAGGACAGCAACTGAAACCCCTCTGGCAATGGAGGCTGGACCTTTGGAAC 307  
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102  
DB 308 GCTCCTTTCACTACCATGCTTTCCACTTCTCAGAACCGCCAGACCTTCATTACCTCAGTC 367  
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTyrGluTyrProGly 122  
DB 368 ATCAAAATTTCTCGTCAGTATGGTGTGATGACTGGACCTGGACTGGGAATACCCAGGC 427  
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142  
DB 428 TCAGCTGGAGCCCTCTCAGACAGCAAGCATCTTTCACCTGCTCTCTGGTGAAGAAATGCT 487  
QY 143 AsnAlaPheGlnGlnGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162  
DB 488 GAAGCTTTTGACGAGGCTATTGAGAGCAACAGCCAGACGATGATGCTTACTGCTGCT 547  
QY 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182  
DB 548 GTAGCTGGTGGGATTTCCAACTATCCAGGCTGGCTATGAGATCCCTGAACTTTCTAAGTAC 607  
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTyrTyrGluLysValThrGly 202  
DB 608 CTGGATTTTCATGCTCATGATCATATGATGATGCTCATGCTCTCTGGAGGCTACACTGG 667  
QY 203 HisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAspLeuAsnVal 222  
DB 668 GAGATAGTCTCTTTACAAATACCTACTGAGACTGGTAGCAATGCCTTACCTCAATGTG 727  
QY 223 AspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGly 242  
DB 728 GATTATGTCATGAACATTATTGGGAAGAACATGAGGCCCCAGCTGAGAACTCATTTGTGGA 787  
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262  
DB 788 TTCCAGAGTAGTGGACACACCTTCATCTCTGAGAAACCCCTCTGATTAATGGAATTTGCTGC 847  
QY 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAlaTyr 282  
DB 848 CCTACTCTGTGTGATGGCCCTGCTGGCCCTATACCAGACAGGCTGGGTCTGGGGCTAC 907  
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
DB 908 TATGAGATTGCACCTTTCTGAGAAGTGAGGACCTGAGGCTCTGGGATGCCCTCCCAAGAA 967  
QY 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspValGluSerPheLys 320  
DB 968 GTGCCCTATGCTTAAAGCCAAAGAGTGCTTGGCTATGACAAATATCAAGAGCTTCAGT 1027  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
DB 1028 GTTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCCCATTCAC 1087  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
DB 1088 CTTGATGACTTCACTGGCTCTTCTGTGATGAGGAAATTTCTCTGACTTCTACTTTG 1147  
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
DB 1148 AACAAAGCCCTGGCATATCCACTGAAGGTTGACAGCTCCTCTGACGTG----- 1195  
QY 381 ProGlyGlnProSerGluPro-----Glu 388

DB 1196 -----CCTCCGAGCCAGTACTACTCTCCAGGAAGTGGGAGTGGAGC 1246  
QY 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro 408  
DB 1247 TCCGGAGGAAGCTCTGGAGGAGTGGATTTGTGCCGCAAGCAGATGGCCTCTACCC 1306  
QY 409 AsnProArgGluArgSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428  
DB 1307 GTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCAT 1366  
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTyr 444  
DB 1367 TGTCAAGCAGGCGCTGTTTGTATACCACTGCTGTAATGCTGCAACTGG 1414  
RESULT 15  
BC034548 1557 bp mRNA linear ROD 07-AUG-2002  
LOCUS Mus musculus, chitinase, acidic, clone MGC:19045 IMAGE:4189080,  
DEFINITION mRNA, complete cds.  
ACCESSION BC034548  
VERSION BC034548.1 GI:21961190  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1557)  
Strausberg,R.  
Direct Submission  
AUTHORS Submitted (24-JUL-2002) National Institutes of Health, Mammalian  
TITLE Gene Collection (MGC), Cancer Genomics Office, National Cancer  
JOURNAL Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 24 Row: h Column: 16  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein.  
FEATURES  
Location/Qualifiers  
1. 1557  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N"  
/clone="MGC:19045 IMAGE:4189080"  
/tissue\_type="Salivary gland, 10 week old female mouse"  
/clone\_lib="NCI CGAP\_SG2"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
4. 1425  
/codon\_start=1  
/product="chitinase, acidic"  
/protein\_id="AAH34548.1"  
/db\_xref="GI:21961191"  
/db\_xref="LocusID:81600"  
/translation="MAKLLVLTGLALLNAQLGSAYNLICYFTNWAQYRPGLSFKPD  
DINPLCLTHLYAFAGMONEITTEWNDVLYKAFNDLKNRNSKLKTLAIGWNFG

TAPFTMTVSTQNRQTFITTSVIKFLRQYFDGLDLDWEYPSGRSGSPQDKHLFTVLVK  
ENREAEQBAISNRRPRLMTVAAGGIGNIOAGYEIPBLSKYLDPIHVMYDLAGSW  
EYGTGNSPLKYKPYETGTGNAYLVNDVVMYKNGCAPAEKLI VGPPEYGHFTILRNP  
SDNGICAPTSQDGPAGPYTRQAGFWAYEICTFLRSGATEVMDASQEVYAYKANEWL  
GDNIKISFSYKQWMLKQNFQGMWIMWALDDFTGFSFCQDQKFPPLSTLUNKALGISTE  
GCTAPDVPBPVPTTPGSGGGSGSGSGSGGSGGFCADKADGLYPVADDRNRFQWCCING  
ITYQHCORAGLVFDTSNCNCNP"

BASE COUNT 410 a 382 c 366 g 399 t

ORIGIN

Alignment Scores:

Pred. No.: 7,01e-92 Length: 1557  
Score: 1309.50 Matches: 236  
Percent Similarity: 67.32% Conservative: 71  
Best Local Similarity: 51.75% Mismatches: 128  
Query Match: 54.61% Indels: 21  
DB: 10 Gaps: 3

US-10-004-219B-10 (1-445) x BC034548 (1-1557)

QY 3 LeuValCysThrPheThrAsnTPAlaGlnTyrArgGlnGlyGluAlaAArgPheLeuPro 22  
Db 73 CTGATATGCTATTTCACCACTGGGCCCGAGTATCGGCCAGGCTCGGGGAGCTTCAAGCCT 132  
QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42  
Db 133 GATGACATTAACCCCTCGCTGTGTACTACCTGATCTATGCTTTCTGGGATGCAGAAC 192  
QY 43 HisGlnLeuSerThrThrGlnTyrAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62  
Db 193 AATGAGATCACCACCATAGAAATGAATGATGTACTCTCTATAAAGCTTTCAATGACTTG 252  
QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr 82  
Db 253 AAAACAGGAACAGCAACTGAAACCCCTCGCAATTGGAGCTGGAACTTTGGAAT 312  
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102  
Db 313 GCTCCTTTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACCTTCAATACCTCAGTC 372  
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGlnTyrProGly 122  
Db 373 ATCAAAATTTCTGCGTCAGTATGGATTTGATGGACTGGACCTGGGAAATACCCAGGC 432  
QY 123 SerGlnGlySerProAlaValAspLysGluAArgPheThrThrLeuValGlnAspLeuAla 142  
Db 433 TCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGCTCTGGTGAAGGAATGCGT 492  
QY 143 AsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162  
Db 493 GAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTTACTGTGCT 552  
QY 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182  
Db 553 GTAGCTGGTGGGATTTCCAAACATCCAGGCTGGGTATGAGATCCCTGAACTTTCTAAGTAC 612  
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGlnLysValThrGly 202  
Db 613 CTGATTTTCACTCATGTCATGATATGACTCTCATGCTCTCGGGAGGGCTACACTGGG 672  
QY 203 HisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222  
Db 673 GAGATAGTCTCTTTTACAAATACCCCTACTGAGACTGGTAGCAATGCCCTACCTCAATGTG 732  
QY 223 AspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGly 242  
Db 733 GATTATGTCTATGAACTATTGGAAAGAACAAATGGAGGCCCGCCAGCTGAGAACTATTGTGGA 792  
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGlyAla 262  
Db 793 TTCCCAGATATGGACACACCTTCACTCTGAGAAACCCCTCTGATTAATGGAATGGTGCC 852  
QY 263 ProAlaThrGlySerGlyThrProGlyThrProPheThrLysGlyGlyMetLeuAlaTyr 282

Search completed: June 29, 2003, 23:53:45

Job time : 2918.5 secs

Db 853 CCTACCTCTGTGTGATGGCCCTGCTGGGCCCTATACACAGAGCTGGGTCTTGGGCCTAC 912  
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 913 TATGAGATTGGACCTTTCTTGAGAAAGTGAGGCCACTGAGGCTCGGATGGCTCCCAAGAA 972  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320  
Db 973 GTGCCCTATGCCATAAGGCCAACGAGTGGCTTGCTGCTATGACAATATCAAGAGCTTCAGT 1032  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340  
Db 1033 GTTAAGGCTCAGTGGCTTAAGCAGAACCAATTTTGGAGGTGCCATGATCTGGGCCATTGAC 1092  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
Db 1093 CTGTATGACTTTCACCTGGCTCTTTCTGTGATCAGGGAATAATTTCTGACTTCTACTTTG 1152  
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
Db 1153 AACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCCTGACGTG----- 1200  
QY 381 ProGlyGlnProSerGluPro-----Glu 388  
Db 1201 -----CCTTCGAGCCAGTGACTACTCTCCAGGAAGTGGAGTGGGGTGGGAAGC 1251  
QY 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro 408  
Db 1252 TCCGGAGGAAGCTCTGGAGGCGAGTGTCTGTGCCGCAAAAGCAGATGGCCTCTACCCCT 1311  
QY 409 AsnProArgGluArgSerSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428  
Db 1312 GTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCAT 1371  
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysThrTrp 444  
Db 1372 TGTCAAGCAGGCGCTTGTGTTTGTATACCAGCTGTATAATGTGCAACTGG 1419



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:02:46 ; Search time 244.223 Seconds  
(without alignments)  
4103.372 Million cell updates/sec

Title: US-10-004-219B-10

Perfect score: 2398

Sequence: 1 AKLVCFYTNWAQRQGEARF.....QQSCPTGLVFSNCKCTWN 445

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xl  
-O=Cgn2\_1/USPO2\_spool/US10004219/runat\_24062003\_160227\_8151/app\_query.fasta\_1.1877  
-DB=N Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CGN 1.1.626 @runat\_24062003\_160227\_8151 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2398	100.0	1636	19	AAV10435	Human chitinase cl
2	2398	100.0	1636	20	AAZ21847	MO-218 clone of hu
3	2398	100.0	1636	22	AAO31759	Human chitinase CD
4	2398	100.0	1636	24	ABL57380	Human chitinase CD
5	2394	99.8	1643	18	AAV50833	Human chitinase CD
6	2392	99.7	1656	19	AAV10436	Human chitinase cl
7	2392	99.7	1656	20	AAZ21848	MO-13B clone of hu
8	2392	99.7	1656	22	AAO31760	Human chitinase CD
9	2392	99.7	1656	24	ABL57381	Human chitinase CD
10	2391	99.7	1768	18	AAV89181	Human chitotriosid
11	2372	98.9	1637	18	AAV89180	Human chitotriosid
12	2348.5	97.9	1713	18	AAV50834	Human chitinase CD
13	1323	55.2	1368	22	AAH42013	Disease treatment
14	1323	55.2	1678	22	AAH42025	Disease treatment
15	1072.5	44.7	1369	24	ABN59993	Novel human coding
16	1061	44.2	1594	19	AAV21688	DNA encoding a hum
17	1053	43.9	1432	19	AAV21689	DNA encoding a hum
18	1053	43.9	1594	19	AAV21687	DNA encoding a hum
19	1046.5	43.6	1433	19	AAV13925	Human cartilage gp
20	1046.5	43.6	1496	18	AAV97127	Human cartilage gl
21	1046.5	43.6	1526	19	AAV13926	Human cartilage gp
22	1035.5	43.2	1152	18	AAV9452	Bovine whey protei
23	1030	43.0	1681	16	AAQ85245	YKL-40 gene. Homo
24	1030	43.0	1681	21	AAZ94901	Human cancer marke
25	1015.5	42.3	2346	23	ABL03597	Drosophila melanog
26	996	41.5	1391	24	AAZ62615	cDNA sequence #402
27	975.5	40.7	17368	23	ABL03596	Drosophila melanog
28	967.5	40.3	1469	22	AAH42023	Disease treatment
29	964.5	40.2	1994	16	AAQ90442	Bovine oviduct spe
30	951	39.7	1474	22	AAQ23078	Osteoarthritis tis
31	942	39.3	2504	16	AAQ90443	Murine oviduct spe
32	922	38.4	2366	16	AAQ90444	Hamster oviduct sp
33	919.5	38.3	13497	23	ABL03557	Drosophila melanog
34	848	35.4	1433	23	ABL15099	Drosophila melanog
35	844.5	35.2	16489	23	ABL03556	Drosophila melanog
36	830	34.6	3552	23	ABL15098	Drosophila melanog
37	824	34.4	2452	18	AAV62557	Manduca sexta larv
38	797	33.2	966	9	AAH81756	Gene encoding poly
39	786.5	32.8	2089	23	ABL29841	Drosophila melanog
40	762	31.8	1470	21	AAZ38589	D. pteronyssius 98
41	762	31.8	1470	21	AAZ38590	D. pteronyssius 98
42	762	31.8	1470	24	ABK69585	Der HMM-map polynu
43	762	31.8	1470	24	ABK69586	Der HMM-map polynu
44	762	31.8	1527	21	AAZ38587	D. pteronyssius 98
45	762	31.8	1527	21	AAZ38588	D. pteronyssius 98

ALIGNMENTS

RESULT 1

AAV10435

ID AAV10435 standard; cDNA; 1636 BP.

XX AAV10435;

AC AAV10435;

XX 15-JUN-1998 (first entry)

DT	DT	DT	DT	DT	DT	Human chitinase clone MO-218 CDNA.
DE	DE	DE	DE	DE	DE	Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
XX	XX	XX	XX	XX	XX	Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW	KW	KW	KW	KW	KW	rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX	XX	XX	XX	XX	XX	Homo sapiens.
OS	OS	OS	OS	OS	OS	Key
PH	PH	PH	PH	PH	PH	Location/Qualifiers
FT	FT	FT	FT	FT	FT	2..1402

FT sig\_peptide /\*tag= a  
 FT 2..64  
 FT mat\_peptide /\*tag= b  
 FT 65..1399  
 FT /\*tag= c  
 FT /product= chitinase  
 FT /note= "from clone MO-218"  
 XX  
 PN W09747752-A1.  
 XX  
 PD 18-DEC-1997.  
 XX  
 XX 16-JUN-1997; 97WO-US10460.  
 XX  
 PR 14-JUN-1996; 96US-0663618.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Gray PW;  
 XX  
 DR WPI; 1998-052316/05.  
 DR P-PSDB; AAW40259.  
 XX  
 XX Nucleic acids encoding human chitinase - useful as antifungal  
 PT agents, especially in combination with other antifungals  
 PT  
 XX Claim 3; Page 38-40; 63pp; English.  
 XX  
 CC This sequence encodes a novel human chitinase isolated from clone MO-218.  
 CC Chitinases are useful for treating or preventing fungal infection and  
 CC as immunogens for generating antibodies which are used to purify, detect  
 CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The  
 CC nucleic acid sequence of the chitinase is also useful as a probe to  
 CC identify and isolate genomic DNA encoding chitinases or similar proteins,  
 CC or cells expressing them or to generate transgenic ('knockout') rodents.  
 CC It can also be used in hybridisation assays and to detect genetic  
 CC alterations in the chitinase gene related to disease. Agents that inhibit  
 CC this protein may be useful in treatment of Gaucher's disease and  
 CC rheumatoid arthritis, where overexpression of the protein can damage  
 CC the extracellular matrix. Chitinase also improves the activity of other  
 CC antifungal agents and may allow a reduction in the dose of such agents,  
 CC and thus of their side effects.  
 XX  
 SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8,43e-201 Length: 1636  
 Score: 2398.00 Matches: 445  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 19 Gaps: 0  
 US-10-004-219B-10 (1-445) x AAV10435 (1-1636)  
 QY 1 AlaLysLeuValCysTyrPheThrAsnTTPAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
 DB 65 GCMAAACTGGTCTGCTTCTTCCAACTGGGCCCATGACAGAGGGGAGGCTCGCTTC 124  
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
 DB 125 CTGCCCAAGGACTTGGACCCAGCCTTTCACCCACCTCATCTAGCCCTTCGCTGCATG 184  
 QY 41 ThrAsnHisGlnLeuSerThrThrGluTTPAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 DB 185 ACCAACCCACGAGTGAACCACTGAGTGGAAATGACGAGACTCTCTACAGGAGTTCAAT 244  
 QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTTPAsnPhe 80  
 DB 245 GGCTTGAAGAGAGTGAATCCCAAGCTGAAGACCTGTAGCCATCGAGGCTTGAATTTC 304  
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnArgGlnThrPheValAsn 100

305 GGCACCTCAGAGTTTCACAGATATGTCAGCCAGCGCAACACCGTCAGACCTTTGTCAAC 364  
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
 DB 365 TCGGCCATCAGGTTTCTCGCAATACAGCTTTGACGGCTTGACCTTGAGTGGAGATAC 424  
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
 DB 425 CCAGGAGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCACAAACCTGGTACAGGAC 484  
 QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
 DB 485 TTGGCCAAATGCTTCAGCAGGAGCCAGACCTCAGGAGGAAACGCGCTTCTTCTAGT 544  
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
 DB 545 GCAGCGGTTCAGCTGGGCAGACCTATGTGGATGCTGGATACGAGGTGGACAAATCGCC 604  
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
 DB 605 CAGAACCTGGATTTTGTCAACCTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTC 664  
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyAlaAlaSerLeu 220  
 DB 665 ACGGACATACAGCCCTCTTACAAGAGGCAAGAGAGAGTGGTGCAGCAGCCAGCCTC 724  
 QY 221 AsnValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIle 240  
 DB 725 AACGTGGATGCTGTGTGCACAGTGGCTGCAAGAGGGAGCCCTGCCAGCAAGCTGATC 784  
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
 DB 785 CTTGGCATGCTTACCTACGAGGCTCTCTCACATGGCTCTCATCAGACACAGAGTG 844  
 QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
 DB 845 GGGGCCCCAGCCACAGGCTCTGGCACTCCAGGCCCTTCACCAAGAGAGGAGGATGCTG 904  
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
 DB 905 GCCTACTATGAAGTCTGCTCTCGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 964  
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
 DB 965 GTGCCCTACATCTTCGGGACACACAGTGGTGGCTTTGATGATGTGGAGAGCTTCAAA 1024  
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
 DB 1025 ACCAAGGTCAGCTATCTGAAGCAGAAGGAGCTGGGGGGGCCCATGGTCTGGGCACTGGAC 1084  
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
 DB 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1144  
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
 DB 1145 CGGCAAGAACTGAGTCTTCCATCTTCCATCTTCCAGCACCCAGAGCTTGAAGTTCCAAA 1204  
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
 DB 1205 CCAGGTCAGCCCTCTGAACCTGAGCATGGCCCGCCAGCCCTGGACAAGACACGCTTCTGCCAG 1264  
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
 DB 1265 GGCAGAGCTGATGGGCTCTATCCCAATCTCTGGAAACGGTCCAGCTTCTACAGCTGTGCA 1324  
 QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
 DB 1325 GCGGGGGGCTGTTCAGCAAGAGCTGCCCGACAGCGCTGGTGTTCAGCAACCTCTCTGCAAA 1384  
 QY 441 CysCysThrTrpAsn 445  
 DB 1385 TGCTGCACCTGGAAT 1399

## RESULT 2

AAZ21847

ID AAZ21847 standard; DNA; 1636 BP.

XX AAZ21847;

XX AAZ21847;

XX 10-DEC-1999 (first entry)

XX MO-218 clone of human Chitinase, with noncoding 5'/3' regions.

XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;  
KW organ transplant; parasite; chitin-binding; allele; vector;  
KW truncated protein; ds.

XX Homo sapiens.

XX Key

FH Location/Qualifiers

CDS 2..1402

FT /\*tag= a

FT /product= Human\_Chitinase

FT sig\_peptide

FT /\*tag= b

FT /product= Signal\_peptide

FT mat\_peptide

FT 66..1402

FT /\*tag= c

FT /product= Mature\_protein

XX W09946390-Al.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Novel chitin-binding fragments of human chitinase used to treat fungal infections in animals

Example 1; Page 55-57; 83pp; English.

This is the nucleotide sequence of an allelic form of the human chitinase enzyme, which is capable of degrading Chitin (a linear homo polymer of beta-1,4-linked N-acetylglucosamine residues). Chitinase fragments can be used to screen for proteins or other molecules that specifically bind to the chitin-binding domain of human chitinase or that modulate its activity. These compounds are useful for immunization, as well as for purifying chitinase, as well as for detection and quantification of chitinase. Polynucleotide fragments of the invention are useful as a source of probes and primers, and to express the proteins recombinantly. The chitinase fragments, when conjugated to antifungal compounds, are used to treat animals, especially humans, infected with chitin-containing parasites such as fungi. Fungal infection treated include candidiasis, aspergillosis, mucormycosis, blastomycosis, paracoccidioidomycosis, coccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, and dermatophytoses. Chitin can be degraded by the enzyme chitinase. Use of whole chitinase protein for treating infections, especially fungal infections, is problematic. In view of the increasing incidents of life-threatening fungal infection in e.g. immunocompromised individuals, there exists a need for identifying new compounds for treating fungal infection. The chitin-binding fragments of the present invention provide this need.

Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

## Alignment Scores:

Pred. No.: 8, 43e-201 Length: 1636  
Score: 2398.00 Matches: 445  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-10-004-219B-10 (1-445) x AAZ21847 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
DB 65 GCAAACTGGTCTGCTACTTACCAACTGGGCCCTAGTACAGACAGGGGAGGCTCGCTTC 124  
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40  
DB 125 CTGCCCAAGGACTTGGACCCCGCCCTTTGCACCCACCTCACTACGCCCTTCGCTGGCATG 184  
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
DB 185 ACCAACCAACAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACCCAGGAGTTCAAT 244  
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaLeuGlyGlyTrpAsnPhe 80  
DB 245 GGCCTGAAAGAGATGANTCCCAAGCTGAAGACCCCTGTAGCCATCGGAGGCTGGAAATTC 304  
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
DB 305 GGCACCTCAGAGTTTCCAGATATGTTAGCCACCGCCACACCCGTACAGACCTTTGTCAAC 364  
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
DB 365 TCGGCCCATCAGGTTTCTGGCAAAATACAGCTTTGACGGCCTTGACTTGGAGTAC 424  
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
DB 425 CCAGGAAGCCAGGGGAGGCCCTGCCGTAGACAGGAGCGCTTCACAAACCTTGTACAGGAC 484  
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
DB 485 TTGGCAATGCTTCCAGAGGAAGCCAGACCTCAGGGAGGAAGCGCTTCTTCTGAGT 544  
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
DB 545 GCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATACGAGGTGGACAAATCGCC 604  
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
DB 605 CAGAACCTGGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTC 664  
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220  
DB 665 ACGGACATTAACAGCCCTCTTACAAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCTTC 724  
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240  
DB 725 AACGTGGATGCTGCTGTCACACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATC 784  
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
DB 785 CTTGGCATGCTTACCTACGAGCGCTCTTACACCTGGCCCTCTCATCAGACACCAAGAGTG 844  
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
DB 845 GGGGGCCCAAGCCAGGAGGCTGGCACTCCAGGCCCTTCCACCAAGAGAGGAGGATGCTG 904  
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
DB 905 GCCTACTATGAGTCTGCTCTCTGGAGGGGGCCCAACACAGAGATCCAGGATCAGAGAG 964  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
DB 965 GTGGCCTACATCTTCCGGGACCAACCAAGTGGTGGGCTTTGTATGATGTGGAGAGCTTCAAA 1024

QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrrAlaLeuAsp 340  
 DB 1025 ACCAAGTCAGCTATCTGAGAGCAGAGGAGCTGGGGGGCCATGTCTGGGCACTGGAC 1084

QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuLeuGlnThrLeu 360  
 DB 1085 TTAGATGACTTTGCGGCTTCTCTGCAACAGGGCGGATACCCCTCATCCAGACGCTA 1144

QY 361 ArgGlnLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
 DB 1145 CGCAGGAAGCTGAGTCTTCATACTTGCCTTCAGGACCCAGAGTTGAAGTTCCAAA 1204

QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
 DB 1205 CAGGTCAGCCCTCTGACCTGAGCATGGCCAGCCCTGGACACACAGCTTCTGCCAG 1264

QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
 DB 1265 GGCAGAGCTGATGGCTCTATCCCAATCTCTCGGACGGTCCAGCTTCTACAGCTGTGCA 1324

QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
 DB 1325 GCGGGCGGCTGTTCAGCAAGCTGCCGACAGGCTGTGTTCAGCAACTCTCCAAA 1384

QY 441 CysCysThrTrrAsn 445  
 DB 1385 TGCTGCACTGGAT 1399

RESULT 3  
 AAD03759  
 ID AAD03759 standard; cDNA; 1636 BP.  
 AC AAD03759;  
 XX  
 DT 19-JUN-2001 (first entry)  
 DE Human chitinase cDNA from clone pMO-218.  
 KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;  
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;  
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatomycoses;  
 KW clone pMO-218; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 2..1402  
 FT /\*tag= a  
 FT /product= "Human chitinase from clone pMO-218"  
 FT sig\_peptide 2..64  
 FT /\*tag= b  
 FT mat\_peptide 65..1399  
 FT /\*tag= c  
 FT /product= "Human mature chitinase from clone pMO-218"  
 XX  
 PN WO200123430-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 28-SEP-2000; 2000WO-US26960.  
 XX  
 PR 30-SEP-1999; 99US-0409918.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;  
 XX  
 DR WPI: 2001-266141/27.  
 DR P-PSDB; AAE00432.  
 XX  
 FT Novel chitinase immunoglobulin fusion product, useful for treating

PT fungal infections and reducing the amount of a non-chitinase antifungal  
 PT agent needed for the treatment -  
 XX  
 PS Claim 2; Page 29-31; 39pp; English.  
 XX  
 CC The present invention relates to a chitinase immunoglobulin (Ig) fusion  
 CC product, comprising a human chitinase fused to at least a portion of an  
 CC immunoglobulin chain. The fusion product is useful for treating fungal  
 CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,  
 CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
 CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatomycoses.  
 CC The fusion protein is useful for reducing the amount of non-chitinase  
 CC antifungal agent needed to exert an antifungal activity. The fusion  
 CC protein is also useful for preparing a medicament for the prophylactic  
 CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin  
 CC fusion product has unexpectedly improved serum half-life and formulation  
 CC properties. The present sequence is human chitinase cDNA from clone  
 CC pMO-218. Chitinase enzyme degrades chitin which is a homopolymer of  
 CC beta-(1,4)-linked N-acetylglucosamine residues.

SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

## Alignment Scores:

Pred. No.: 8,43e-201 Length: 1636  
 Score: 2398.00 Matches: 445  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-10-004-219B-10 (1-445) x AAD03759 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrrAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
 DB 65 GCAAACTCGTCTGCTACTTCCAACTGGGCCCACTAGACAGAGGGAGGCTCGCTTC 124

QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40  
 DB 125 CTGCCCAAGGACTTGGACCCCGACGCTTTGACCCACCTCATCTACGCTTCGCTGGCATG 184

QY 41 ThrAsnHisGlnLeuSerThrThrGluTrrAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 DB 185 ACCAACCAACCACTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGAGGTTCAAT 244

QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrrAsnPhe 80  
 DB 245 GGCCTGAGAAGATGATCCCAAGCTGAGACCCCTGTAGCCATCGAGGCTGGAAATTC 304

QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
 DB 305 GGCACCTCAGAGTTCCACAGATATGTCAGCAGCGGCCAACACCCGTCAGACCTTTGTCAAC 364

QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrrPduLys 120  
 DB 365 TCGGCCATCAGGTTTCTGCGCAATACAGCTTTCAGCGCTTACCTTACCTGGAGGATC 424

QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
 DB 425 CCAGGAAGCCAGGAGGCGCTGCTGCTAGACAGGAGCGCTTCACAACTGCTGTGACAGGAC 484

QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
 DB 485 TTGGCCAATGCTTCCAGCAGGAGGCCAGACCTCAGGGAAGGAACGCTTCTCTGAGT 544

QY 161 AlaAlaValProAlaGlyGlnThrThrTrrValAspAlaGlyTyrGluValAspLysIleAla 180  
 DB 545 GCAGCGGTTCCAGCTGGGCGAGACCTTATGTGATGCTGATACGAGTGGACAAAATCGCC 604

QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTrrAspPheHisGlySerTrrPduLysVal 200  
 DB 605 CAGAACCTTGGATTTGTCAACCTTATGCTTACGACTTCCATGGCTCTCTGGAGAGGTC 664

QY 201 ThrGlyHisAsnSerProLeuTrrLysArgGlnGluSerGlyAlaAlaSerLeu 220

Db 665 ACAGGACATAACAGCCCTCTACAGAGGCAAGAGAGTGGTGCAGCAGCCAGCCTC 724  
Qy 221 AenValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuile 240  
Db 725 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGACCCCTGCCAGCAAGCTGATC 784  
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
Db 785 CTTGGCATGCCCTACCTACGGACGCTCTTCACACTGGCTCTCTCATCAGACACAGAGTG 844  
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlnGlyGlyMetLeu 280  
Db 845 GGGGCCCCAGCCACAGGGTCTGCACTCCAGGCCCTTCACCAAGAGAGGGATGCTG 904  
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 905 GCCTACTAATGAAGTCTGCTCTCGGAAGGGGGCCACCAACAGAGAAATCCAGGATCAGAG 964  
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
Db 965 GTGCCCTACATCTTCGGGACACACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAA 1024  
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
Db 1025 ACCAAGGTCACTATCTGAAGCAGAGAGGACTGGCGGGGCCATGCTCTGGGCACTGGAC 1084  
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
Db 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGGGCCGATACCCCTCATCCAGAGCTTA 1144  
Qy 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
Db 1145 CGGCAGGAAGTCACTTCCATCTTCCAGGACCCAGAGCTTGAAGTTCCAAA 1204  
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
Db 1205 CCAGGTCAAGCTCTGCACTGAGATGAGCCCGCCAGCCCTGGCAAGACACGCTTCTGCCAG 1264  
Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerPheTyrSerCysAla 420  
Db 1265 GCAAGAGTGTATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCA 1324  
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
Db 1325 GGGGGGGGCTTTCCAGAAAGCTGCCGACAGGCTGGTGTTCAGCACTCTCTGCAAA 1384  
Qy 441 CysCysThrTrpAsn 445  
Db 1385 TGCTGCACCTGGAAT 1399

## RESULT 4

ABL57380  
ID ABL57380 standard; cDNA; 1636 BP.

AC ABL57380;

DT 12-AUG-2002 (first entry)

DE Human chitinase cDNA clone MO-218.

Chitinase; enzyme; human; fungicide; antifungal; infection;  
KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;  
KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
KW Pneumocystis; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers  
FH 2..1402  
FT CDS

FT /\*tag= a  
/product= "Chitinase"

sig\_peptide 2..64  
mat\_peptide /\*tag= b  
65..1399  
/\*tag= c  
US6372212-B1.  
16-APR-2002.  
16-JUN-1997; 97US-0877599.  
14-JUN-1996; 96US-0663618.  
(ICOS-) ICOS CORP.  
Gray PW;  
WPI; 2002-442449/47.  
P-PSDB; ABB76291.

Co-administering chitinase to improve the effectiveness of fungicidal drugs e.g. amphotericin B or itraconazole, useful for treating fungal infections e.g. candidiasis, coccidioidomycosis and blastomycosis -  
Example 1; Column 23-28; 26pp; English.

The present sequence is that of cDNA clone MO-218 (ATCC 98077) encoding human chitinase (see ABB76291). The clone was isolated from a cDNA library prepared from peripheral blood monocyte-derived macrophages following sequence analysis. Also isolated was clone MO-13B (see ABL57381), which contains a single nucleotide difference in the coding region, changing the encoded amino acid at position 81 of the mature protein from glycine to serine. Northern blots showed highest chitinase gene expression in lung and ovary tissues. Expression in lung is consistent with a protective role against pathogenic organisms that contain chitin. The invention provides human chitinase polynucleotides and polypeptides, and materials and methods for the recombinant production of human chitinase products, which are expected to be useful as products for treating fungal infections or for the development of such products. Human chitinase has a synergistic effect on the actions of other fungicides. It can be administered to improve the antifungal activity of a non-chitinase antifungal agent, especially amphotericin B or itraconazole, in the treatment of a fungal infection such as candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis, dermatophytoses and Pneumocystis infections (all claimed). In particular, the fungal infection involves Candida, Aspergillus and/or Cryptococcus spp., whose growth is not effectively inhibited by contact with human chitinase alone.

SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

## Alignment Scores:

Pred. No.: 8, 43e-201 Length: 1636  
Score: 2398.00 Matches: 445  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-10-004-219B-10 (1-445) x ABL57380 (1-1636)

Qy 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyAlaArgPhe 20  
Db 65 GCAAAACTGGTCTGCTACTTCCAACTGGGCCCATAGTACAGACAGGGGAGCTCGCTTC 124  
Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 125 CTGCCCAAGGAGCTTGGACCCCGAGCCCTTTGGACCCCACTCATCTACGCTTCGCTGGCATG 184  
Qy 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60

```
Db 185 ACCAACCCAGCTGAGCACCCTGAGTGAATGACGAGACTCTTACCCAGGAGTTCAAT 244
Qy 61 GlyLeuLysMetAsnProLysLeuLysLeuLeuAlaLeuGlyGlyTyrAsnPhe 80
Db 245 GGCCTGAAGAAGATGAATCCAGCTGAGAGCCCTGTGACCATCGAGGCTGGAAATTC 304
Qy 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db 305 GGCACCTCAGAGTTTCAGATATGTTAGCCACCGCCCAACCAACCGTCAGACCTTTGTCAAC 364
Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTyrGluTyr 120
Db 365 TCGGCCATCAGGTTCTGGCCAAATACAGCTTTGACGGCCTTGACCTGACCTGGAGTAC 424
Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
Db 425 CCAGGAAGCAGGGAGCCCTGCGGTAGACAGAGCGCTTCACACCCCTGGTACAGGAC 484
Qy 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
Db 485 TTGGCCCAATGCCCTTCAGCAGGAAGCCAGACCTCAGGGAAGAACGCTTCTTCTGAGT 544
Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
Db 545 GCAGCGGTTCCAGCTGGCAGACCTATGTGATGCTGATACGAGGTGGACAAAATCGCC 604
Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHicGlySerTyrGluLysVal 200
Db 605 CAGAACCTGGATTTGTCAACCTTATGGCTACGACTTCATGGCTCTTGGGAAGAGGTC 664
Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaLaserLeu 220
Db 665 ACGGACATACAGCCCTCTACAGAGGCAAGAGAGTGTGTCAGCAGCAGCCCTC 724
Qy 221 AsnValAspAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIle 240
Db 725 AACGTGGATGCTGCTGCAACAGTGGCTGCAAGAGGAGCCCTGCCAGCAAGCTGATC 784
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db 785 CTTGGCATGCTACCTACGAGACCTCTTCACTGGCTCTCTCATCAGACACAGAGTG 844
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
Db 845 GGGGCCCCAGCCAGGCTCTGCACTCCAGGCCCCCTTCAACAGGAGGAGGATGCTG 904
Qy 281 AlaTyrTyrGluValCysSerTyrLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 905 GCCTACTATGAGTCTGCTCTGGAAGGGGGCCCAACCAAGAGATCCAGGATCAGAAG 964
Qy 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspValGluSerPheLys 320
Db 965 GTGCCCTACATCTCCGGGCAACACAGTGGTGGCTTTGATGATGGAGAGCTTCAAA 1024
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTyrPalaLeuAsp 340
Db 1025 ACCAAGTTCAGCTATCTGAAGCAGAGAGGACTGGCGGGGCCATGCTGGGCACTGGAC 1084
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1085 TTAGATGACTTTGCGGGCTCTCTGCAACACAGGGCCGATACCCCTCATCCAGAGCTA 1144
Qy 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1145 CGGCAGAACTGAGTCTTCATACTTGGCTTACGACCCCGAGAGCTTGAAGTTCCAAA 1204
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1205 CCAGGTGAGCCCTCTGAACCTGAGCATGGCCCGCCAGCCCTGGCAAGACAGCTTCTGCCAG 1264
Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
```

```
Db 1265 GCGAAAGCTGATGGCTCTATCCCAATCTCCGGAACGGTCCAGCTTTCACAGCTGTGCA 1324
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db 1325 GCGGGCGGCTGTTTCAGCAAGCTGCCGACAGCGCTGGTGTTCAGCAACTCTCTGCAAA 1384
Qy 441 CysCysThrTyrAsn 445
Db 1385 TGCTGCACCTGGAAT 1399

RESULT 5
AAT50833
ID AAT50833 standard; cDNA; 1643 BP.
XX
AC AAT50833;
XX
DT 24-MAR-1997 (first entry)
XX
DE Human chitinase cDNA clone chi.50.
XX
KW Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipodosis; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 13..1413
FT sig_peptide /*tag= a
FT mat_peptide /*tag= b
FT /*tag= c
XX
PN WO9640940-A2.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-NL00225.
XX
PR 07-JUN-1995; 95US-0486839.
XX
PA (UNAM ) UNIV AMSTERDAM.
XX
PI Aerts JWFG;
XX
XX WPI; 1997-118698/11.
DR P-PSDB; AAW08584.
XX
PT New human chitinase - used to treat or prevent infection by
PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
PT foods, implants etc.
PS Claim 2; Page 39-40; 58pp; English.
XX
XX A cDNA clone (AAT50833), designated chi.50, codes for a 50 kDa human
CC chitinase (AAW08584) that is stable to many proteases, active at pH 3-
CC 8 and up to 50 deg, and stable in the circulation. Clones chi.50
CC and chi.39 (see also AAT50834) were isolated from a human macrophage
CC cDNA library using as probe a partial clone obtd. using primers
CC (see also AAT50835-36) based on a chitotriosidase purified from a
CC type 1 Gaucher disease patient. The 2 cDNA clones are the result
CC of alternative splicing of RNA. Chitinase nucleic acid can be used
CC for large-scale prodn. of recombinant human chitinases, or can be
CC incorporated into a gene therapy vector to treat or prevent
CC infection by chitin-contg. pathogens.
SQ Sequence 1643 BP; 364 A; 490 C; 442 G; 347 T; 0 other;

Alignment Scores: 1.9e-200 Length: 1643
Pred. No.: 2394.00 Matches: 444
Score:
```

Percent Similarity: 100.00%    Conservative: 0  
Best Local Similarity: 100.00%    Mismatches: 0  
Query Match: 99.83%    Indels: 0  
DB: 18    Gaps: 0

US-10-004-219B-10 (1-445) x AAT50833 (1-1643)

2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
79 AAACCTGGTCTGCTACTTCCACCACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTCCTG 138  
22 ProLysAspLeuAspProSerLeuCysThrHisLeulleTyrAlaPheAlaGlyMetThr 41  
139 CCCAAGGACTTGGACCCAGCCCTTTGGACCCACCTCATCTAGCCCTTCGCTGGCATGACC 198  
42 AsnHisGlnLeuSerThrTrpGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
199 AACCCACAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGAGATTCAATGCG 258  
62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
259 CTGAAGAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTCGCG 318  
82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
319 ACTCAAGAGTTACAGATATGTTAGCCAGCGCCCAACACCGTCAGACCTTTGTCAACTCG 378  
102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTrpPro 121  
379 GCCATCAGGTTTCTCGGCAAAATACAGCTTTGACGGCTTGACCTTGAGCTGGGAGTACCCA 438  
122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
439 GGAACCCAGGGGAGCCCTGCGGTAGCAAGAGAGCGCTTCACAAACCTGGTACAGACTTG 498  
142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAla 161  
499 GCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAAACGCTTCTTCTGAGTGCA 558  
162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
559 GCGGTTCAGCTGGGCGAGACCTATGTGGATGCTGGATAGAGTGGACAAAATCGCCAG 618  
182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201  
619 AACCTGGATTTGTCAACCTATATGCGCTACGACTTCCATGGCTCTTGGGAGAGGTCAAG 678  
202 GlyHisAsnSerProLeuTyrLysArgGlnGluLysSerGlyAlaAlaAspSerLeuAsn 221  
679 GGACATAACAGCCCTCTTACAAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCCTCAAC 738  
222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeulleLeu 241  
739 GTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGGAGCCCTGCCAGCAAGCTGATCCTT 798  
242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261  
799 GGCATGCTTACCTACGAGCGCTCCTTCACTGGCTCTCTCATCAGACACAGAGTGGGG 858  
262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281  
859 GCCCCAGGCCACAGGGTCTGGCATCTCAGGCCCTTCCACCAAGAGAGGGATGCTGGCC 918  
282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301  
919 TACTATGAAGTCTGCTCTCGGAAGGGGGCCCAACACAGAGAATCCAGGATCAGAGGTG 978  
302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThr 321  
979 CCTACATCTTCGGGACCAACCAAGTGGGTGGCTTTGATGATGGGAGAGCTTCAAAACC 1038  
322 LysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341

1039 AAGGTACGCTATCTCTGAAGACGAAAGGAGCTGGGCGGGGCCATGGTCTGGGCACTGGACTTA 1098  
342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeulleGlnThrLeuArg 361  
1099 GATGACTTTGGCGGCTTCTCTCTGCAACGAGGGCGGATACCCCTCATCCAGACGGTACGG 1158  
362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381  
1159 CAGGAACCTGAGTCTTCCATACCTTGCCTTCAGGCACCCCGAGAGCTTGAAGTTCCAAAACCA 1218  
382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401  
1219 GGTGAGCCCTCTGAACCTGAGCATGGCCCGAGCCCTGGCAAGACACCTTCTGCCAGGCG 1278  
402 LysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAlaAla 421  
1279 AAAGCTGATGGGCTCTATCCCAATCTCGGAGACGGTCCAGCTTCTACAGCTGTGCAGG 1338  
422 GlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441  
1339 GGGCGGCTGTTCCAGCAAGCTGCCGACAGGCTGGTGTTCAGCAACTCCTCTGCAATGC 1398  
442 CysThrTrpAsn 445  
1399 TGCACCTGGAAT 1410

RESULT 6  
AAV10436  
ID AAV10436 standard; cDNA; 1656 BP.  
XX  
AC AAV10436;  
XX  
DT 15-JUN-1998 (first entry)  
XX  
DE Human chitinase clone MO-13B cDNA.  
XX  
Chitinase; human; fungal infection; immunogen; diagnosis; treatment;  
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;  
KW rheumatoid arthritis; overexpression; extracellular matrix; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 27..1427  
FT /\*tag= a  
FT sig\_peptide 27..89  
FT /\*tag= b  
FT mat\_peptide 90..1424  
FT /\*tag= c  
FT /product= chitinase  
FT /note= "from clone MO-13B"  
XX  
XX WO9747752-A1.  
XX  
XX 18-DEC-1997.  
XX  
XX 16-JUN-1997; 97WO-US10460.  
XX  
XX 14-JUN-1996; 96US-0663618.  
XX  
XX (ICOS-) ICOS CORP.  
XX  
XX Gray PW;  
XX  
XX WPI; 1998-052316/05.  
XX  
XX P-PSDB; AAW40260.  
XX  
XX Nucleic acids encoding human chitinase - useful as antifungal  
XX agents, especially in combination with other antifungals  
XX  
XX Claim 9; Page 42-44; 63pp; English.  
XX  
XX This sequence encodes a novel human chitinase isolated from clone MO-13B.

CC Chitinases are useful for treating or preventing fungal infection and  
 CC as immunogens for generating antibodies which are used to purify, detect  
 CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The  
 CC nucleic acid sequence of the chitinase is also useful as a probe to  
 CC identify and isolate genomic DNA encoding chitinases or similar proteins,  
 CC or cells expressing them or to generate transgenic ('knockout') rodents.  
 CC It can also be used in hybridisation assays and to detect genetic  
 CC alterations in the chitinase gene related to disease. Agents that inhibit  
 CC this protein may be useful in treatment of Gaucher's disease and  
 CC rheumatoid arthritis, where overexpression of the protein can damage  
 CC the extracellular matrix. Chitinase also improves the activity of other  
 CC antifungal agents and may allow a reduction in the dose of such agents,  
 CC and thus of their side effects.

XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;  
 SQ

## Alignment Scores:

Pred. No.: 2,886-200 Length: 1656  
 Score: 2392.00 Matches: 444  
 Percent Similarity: 99.78% Conservative: 0  
 Best Local Similarity: 99.78% Mismatches: 1  
 Query Match: 99.75% Indels: 0  
 DB: 19 Gaps: 0

US-10-004-219B-10 (1-445) x AAV10436 (1-1656)

QY 1 AlaValLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyAlaAArgPhe 20  
 DB 90 GCAAAACTGCTGCTGCTACCTTCCCACTGGGCCAGTACAGAGGGGAGGCTCGCTTC 149  
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
 DB 150 CTGCCCAAGGACTTGGACCCACCTTTGCACCCACCTCATCTACGCTTCGCTGGCATG 209  
 QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 DB 210 ACCAACCCAGCTGAGCACCCTAGTGGTGAATGACGAGACTCTTACCAGGAGTTCAAT 269  
 QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80  
 DB 270 GGCTTGAAGAAGATGAATCCAGCTGAGAGCCCTGTAGCCATCGAGGCTGGAAATTC 329  
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
 DB 330 AGCACTCAGAAGTTACAGATATGTTAGCCAGCCCAACACCGTCAGACCTTTGTCAAC 389  
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
 DB 390 TCGGCCATCAGGTTTCTGGCGCAATACAGCTTTGACGGGCTTGACCTTGACCTGGGAGTAC 449  
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
 DB 450 CAGGAAGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCAACACCTGGTACAGGAC 509  
 QY 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160  
 DB 510 TTGGCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAAACGCTTCTCTGAGT 569  
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
 DB 570 GCAGCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACGAGGTGACAAAATCGCC 629  
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
 DB 630 CAGAACCTGATTTGTGCAACCTTATGGCTACGACTTCATGGCTTCTGGGAGAGGTC 689  
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGlnGlnSerGlyAlaAlaAsnLeu 220  
 DB 690 ACGGACATAACAGCCCTCTTACAAGAGCGCAAGAGAGAGTGTGCAGCAGCAGCCCTC 749  
 QY 221 AsnValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProLysLysLeuIle 240  
 DB 750 AACGTGGATGCTGCTGTGAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATC 809

QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
 DB 810 CTTGGCATGCTTACCTAGCGAGCTCTCTTCACTGGCTCTCTCATCAGACACAGAGTG 869  
 QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
 DB 870 GGGGCCCCCAGGCACAGGCTTGGCACTCCAGGCCCTTCAACCAAGGAAGGAGGATGCTG 929  
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
 DB 930 GCCTACTATGAAGTCTGCTCTGGAAGGGGGCCCAACACAGAGAATCCAGATCAGAAG 989  
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320  
 DB 990 GTGCCCTACATCTTCCGGGACCAACAGTGGGTGGCTTGTATGATGTGGAGAGCTTCAAA 1049  
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
 DB 1050 ACCAAGGTGACGTATCTGAAGCAGAAGGAGCTGGGGGGGGCCATGGTCTGGGCACTGGAC 1109  
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
 DB 1110 TTAGATGACTTTGCCGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCCAGACGCTA 1169  
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
 DB 1170 CGGCAAGGAACTGAGTCTTCCATCTTGGCTTCCAGCACCCAGAGCTTGAAGTTCCAAA 1229  
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
 DB 1230 CCAGGTTCAGCCCTCTGAACCTGAGCATGGCCCGCCCTGGACAAGACACGTTCTGCCAG 1289  
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
 DB 1290 GGCAAGCTGATGGGCTTATCCCAATCTCTGGAAACGGTCCAGCTTCTACAGCTGTGCA 1349  
 QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
 DB 1350 GCGGGGCGGCTGTTCAGCAAGCTGCCCGACAGGCTGGTGTTCAGCACTCTTCCAGCA 1409  
 QY 441 CysCysThrTrpAsn 445  
 DB 1410 TGCTGCACCTGGAAT 1424

RESULT 7  
 AAZ21848  
 ID AAZ21848 standard; DNA; 1656 BP.  
 XX AC AAZ21848;  
 XX 10-DEC-1999 (first entry)  
 DT DE MO-13B clone of human Chitinase, with noncoding 5'/3' regions.  
 XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;  
 KW organ transplant; parasite; chitin-binding; allele; vector;  
 KW truncated protein; ds.  
 XX OS Homo sapiens.  
 FT Key Location/Qualifiers  
 FT CDS 27..1427  
 FT /tag= a  
 FT /product= Human\_Chitinase  
 FT 27..89  
 FT /tag= b  
 FT /note= "Signal peptide"  
 FT 90..1427  
 FT /tag= c  
 FT /note= "Mature peptide"  
 XX PN WO9946390-A1.



XX 16-SEP-1999.  
 PD  
 XX 12-MAR-1999; 99WO-US05343.  
 PF  
 XX 12-MAR-1998; 98US-0039198.  
 PR  
 XX (ICOS-) ICOS CORP.  
 PA  
 XX Gray PW, Tjoelker LW;  
 PI  
 XX WPI; 1999-551417/46.  
 DR  
 XX P-PSDB; AAY42426.  
 DR  
 XX Novel chitin-binding fragments of human chitinase used to treat fungal  
 PT infections in animals  
 PT  
 XX Example 1; Page 59-62; 83pp; English.  
 PS  
 XX This is the nucleotide sequence of an allelic form of the human  
 CC chitinase enzyme, which is capable of degrading Chitin (a linear  
 CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).  
 CC Chitinase fragments can be used to screen for proteins or other  
 CC molecules that specifically bind to the chitin-binding domain of human  
 CC chitinase or that modulate its activity. These compounds are useful for  
 CC immunization, as well as for purifying chitinase, as well as for  
 CC detection and quantification of chitinase. Polynucleotide fragments of  
 CC the invention are useful as a source of probes and primers, and to  
 CC express the proteins recombinantly. The chitinase fragments, when  
 CC conjugated to antifungal compounds, are used to treat animals,  
 CC especially humans, infected with chitin-containing parasites such as  
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
 CC sporotrichosis, and dermatophytoses.  
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
 CC protein for treating infections, especially fungal infections, is  
 CC problematic. In view of the increasing incidents of life-threatening  
 CC fungal infection in e.g. immunocompromised individuals, there exists a  
 CC need for identifying new compounds for treating fungal infection. The  
 CC chitin-binding fragments of the present invention provide this need.  
 XX  
 SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:  
 Pred. No.: 2, 88e-200 Length: 1656  
 Score: 2392.00 Matches: 444  
 Percent Similarity: 99.78% Conservative: 0  
 Best Local Similarity: 99.78% Mismatches: 1  
 Query Match: 99.75% Indels: 0  
 DB: 20 Gaps: 0

US-10-004-219B-10 (1-445) x AA221848 (1-1656)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyAlaLahargPhe 20  
 DB 90 GCAAAACTGGTCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 149  
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40  
 DB 150 CTGCCCAAGGACTTGGACCCCAAGCCCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATG 209  
 QY 41 ThrAsnHisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 DB 210 ACCAACCCACCTGAGCACCACTGAGTGGGAATGACGAGACTCTTACCAAGAGTTCAAT 269  
 QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPhe 80  
 DB 270 GSCCTGAAGAAGATGAATCCCAAGCTGAGACCCCTGTTAGCCATCGGAGGCTGGAATTC 329  
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
 DB 330 AGCACTCAGAAGTTCACAGATATGGTAGCCACCGGCCAACACACCGTCAGACCTTTGTCAAC 389

QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
 DB 390 TCGCCCATCAGTTTCTGGCAATACAGCTTTGAGGCTTACCTTGGAGTAC 449  
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
 DB 450 CCAGGAAGCCAGGGGAGCCCTGCGCTAGACAAGAGGCGCTTACAAACCCCTGGTACAGGAC 509  
 QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
 DB 510 TTGCCCAATGCTTCCAGCAGGAAGCCCGACCTTCAGGAAGGAACGCTTCTTCTGAGT 569  
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
 DB 570 GCAGGGTTCAGCTGGGCAGACCTATGTGATGCTGATACGAGGTGACAAATGCC 629  
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
 DB 630 CAGAACCTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGAGAAGTTC 689  
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220  
 DB 690 ACGGACACATAACAGCCCTCTACAAGAGGCAAGAGAGTGGTGCAGCCAGCCCTC 749  
 QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 240  
 DB 750 AACGTGATGCTGTGTGCACAGCTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATC 809  
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
 DB 810 CTTGGCATGCTTACCTACGGAGCGCTCTTCACTGGCTCTCTCATCAGACACCAAGAGTG 869  
 QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280  
 DB 870 GGGGCCCCAGCCACAGGGTCTGGCACTCCAGGCCCTTCCAGAGGAGGAGGATGCTG 929  
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
 DB 930 GCCTACTATGAAGTCTGCTCTCGAAGGGGGGCCCAACAGAGAAATCCAGGATCAGAAG 989  
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
 DB 990 GTGCCCTACATCTTCCGGCACAACCCAGTGGGCTTGTGATGTGGAGAGCTTCAAA 1049  
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340  
 DB 1050 ACCAAGGTCACTATCTGAAGCAGAGGGGACTGGGGCGGGGCCATGGTCTGGGCATGGAC 1109  
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
 DB 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACCCAGGGCCGATACCCCTCATCCAGAGCTA 1169  
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
 DB 1170 CGGCAGGAACCTGAGTCTTCCATATCTTGGTTCAGGCACCCAGGAGCTTGAAGTTCACAAA 1229  
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
 DB 1230 CCAGGTGACGCCCTCTGAACCTGAGATGGGCCCGCCAGCTGGCAAGACACGCTTCTGGCAG 1289  
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
 DB 1290 GGCAAGCTGATGGCTCTATCCCAATCTCTGGGACCGTCCAGCTTCTACAGCTGTGCA 1349  
 QY 421 AlaGlyArgLeuPheGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
 DB 1350 GCGGGGGGCTGTTCCAGAAAGCTGCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA 1409  
 QY 441 CysCysThrTrpAsn 445  
 DB 1410 TGCTGCACCTGGAAT 1424

## RESULT 8

AAD03760  
 ID AAD03760 standard; cDNA; 1656 BP.  
 XX  
 AC AAD03760;  
 XX  
 DT 19-JUN-2001 (first entry)  
 XX  
 DE Human chitinase cDNA from clone pMO-13B.  
 XX  
 KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;  
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;  
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;  
 KW clone pMO-13B; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 27..1427  
 FT /\*tag= a  
 FT /product= "Human chitinase from clone pMO-13B"  
 FT sig\_peptide 27..89  
 FT /\*tag= b  
 FT mat\_peptide 90..1424  
 FT /\*tag= c  
 FT /product= "Human mature chitinase from clone pMO-13B"  
 XX  
 PN WO200123430-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 28-SEP-2000; 2000WO-US26960.  
 XX  
 PR 30-SEP-1999; 99US-0409918.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;  
 XX  
 DR WPI; 2001-266141/27.  
 XX  
 DR P-PSDB; AAE00433.  
 XX  
 PT Novel chitinase immunoglobulin fusion product, useful for treating  
 PT fungal infections and reducing the amount of a non-chitinase antifungal  
 PT agent needed for the treatment  
 XX  
 PS Claim 2; Page 34-36; 39pp; English.  
 XX  
 CC The present invention relates to a chitinase immunoglobulin (Ig) fusion  
 CC product, comprising a human chitinase fused to at least a portion of an  
 CC immunoglobulin chain. The fusion product is useful for treating fungal  
 CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,  
 CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
 CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.  
 CC The fusion protein is useful for reducing the amount of non-chitinase  
 CC antifungal agent needed to exert an antifungal activity. The fusion  
 CC protein is also useful for preparing a medicament for the prophylactic  
 CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin  
 CC fusion product has unexpectedly improved serum half-life and formulation  
 CC properties. The present sequence is human chitinase cDNA from clone  
 CC pMO-13B. Chitinase enzyme degrades chitin which is a homopolymer of  
 CC beta-(1,4)-linked N-acetylglucosamine residues.  
 XX  
 SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

## Alignment Scores:

Pred. No.: 2,88e-200 Length: 1656  
 Score: 2392.00 Matches: 444  
 Percent Similarity: 99.78% Conservative: 0  
 Best Local Similarity: 99.78% Mismatches: 1  
 Query Match: 99.75% Indels: 0  
 DB: 22 Gaps: 0

US-10-004-219B-10 (1-445) x AAD03760 (1-1656)  
 QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
 DB 90 GCAAAACTGGTCTGCTACTTACCACAACTGGGCCCACTAGACAGAGGGAGGCTCGCTTC 149  
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40  
 DB 150 CTGCCCAAGGACTTGGACCCCGAGCCCTTTGCACCCACCTCATCTACGCTTCGCTGGCATG 209  
 QY 41 ThrAsnHisGlnLeuSerThrTrpGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 DB 210 ACCAACCCACCTGAGTGGTGAATGACGAGACTCTCTACGAGGAGTTCAAT 269  
 QY 61 GlyLeuLysLeuMetAsnProLysLeuLysThrLeuLeuAlaLeuGlyGlyTrpAsnPhe 80  
 DB 270 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCTTTAGCCATCGAGGCTGGGAATTTTC 329  
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
 DB 330 AGCACTCAGAACTTCACAGATATGTTAGCCACGCGCAACACCGTCAGACCTTTGTCAAC 389  
 QY 101 SerAlaLeuArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
 DB 390 TCGGGCCATCAGGTTTCTGCGCAAAATACAGCTTTGACGGCTTGACCTTACCTGGGAGTAC 449  
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
 DB 450 CCAGGAAGCCAGGGAGCCCTGCGCTAGCAAGGAGCGCTTCACAACTTGTGTGACAGGAC 509  
 QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
 DB 510 TTGGCCAAATGCTTCCAGCAGGAGAACCCAGACCTCAGGGAGGAACGCTTCTCTGAGT 569  
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysLeuAla 180  
 DB 570 GCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGTGGATAGCAGGTGGACAAAATCGCC 629  
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
 DB 630 CAGAACCTTGGATTTTGTCACTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTC 689  
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu 220  
 DB 690 ACGGCACATAACAGCCCTCTTACAAGAGGCAAGAGAGAGTGGTGCAGACCCAGCCTC 749  
 QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 240  
 DB 750 AACGTGGATGCTGTGTGCAACAGTGGCTGCAGAAAGGGGACCCCTGCCAGCAAGCTGATC 809  
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
 DB 810 CTGTGCATGCCCTACTACGACGCTTCTTCACTTGGCTCTCTCATCAGACACACAGAGTG 869  
 QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280  
 DB 870 GGGGGCCCGAGCCACAGGCTTGGCACTCCAGGCCCTTCCACCAAGGAAGAGGGATGCTG 929  
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
 DB 930 GCCTACTATGAAGTCTGCTCTGGAAGGGGGGCCACCAACAGAGAAATCCAGATCAGAG 989  
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320  
 DB 990 GTGCCCTTACATCTTCCGGGACACACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAA 1049  
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
 DB 1050 ACCAAGGTGAGTATCTGAAGCAGAAAGGAGCTGGGGCGGGGCCATGGTCTGGGCACTGGAC 1109  
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360

Db 1110 TTAGATGACTTTGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGCTTA 1169  
 Qy 361 AtgGlnGluSerLeuProTyrLeuProSerGlyThrProGluLeuValProLys 380  
 Db 1170 CGGAGGAACTGAGCTTCCATCTTCCCTTCCAGCACCCTGAGCTTGAAGTTCAAAA 1229  
 Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
 Db 1230 CCAGGTGAGCCCTCTGAACCTGAGCATGCGCCAGCCCTGGACAGACACGTTCTGCCAG 1289  
 Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
 Db 1290 GGCAAGCTGATGGGCTCTATCCCAATCTCCGGAACGGTCCAGCTTCTACAGCTGTGCA 1349  
 Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
 Db 1350 GGGGGGGGCTGTTCCAGAAAGCTGCCGACAGCGCTGGTGTTCAGCAACTCTTCGAAA 1409  
 Qy 441 CysCysThrTrpAsn 445  
 Db 1410 TGCTGCACCTGGAAT 1424

## RESULT 9

ABL57381  
 ID ABL57381 standard; cDNA; 1656 BP.

AC ABL57381;  
 DT 12-AUG-2002 (first entry)  
 XX Human chitinase cDNA clone MO-13B.

DE Chitinase; enzyme; human; fungicide; antifungal; infection;  
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;  
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
 KW Pneumocystis; gene; ss.  
 OS Homo sapiens.

FH Key Location/Qualifiers  
 FT CDS 27..1427  
 FT /\*tag= a  
 FT /product= "Chitinase"  
 FT sig\_peptide 27..89  
 FT /\*tag= b  
 FT mat\_peptide 90..1424  
 FT /\*tag= c

US6372212-B1.

16-APR-2002.

16-JUN-1997; 97US-0877599.

14-JUN-1996; 96US-0663618.

(ICOS-) ICOS CORP.

Gray FW;

WPI; 2002-442449/47.

P-PSDB; ABB76292.

XX Co-administering chitinase to improve the effectiveness of fungicidal  
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -  
 XX Example 1; Column 31-34; 26pp; English.

CC The present sequence is that of cDNA clone MO-13B (ATCC 98078)

CC encoding human chitinase (see ABB76292). The clone was isolated  
 CC from a cDNA library prepared from peripheral blood monocyte-derived

CC macrophages using a probe based on human chitinase clone MO-218  
 CC (see ABL57380). The 2 clones contain a single nucleotide  
 CC difference in the coding region, causing amino acid position 81 of  
 CC the mature protein to be either serine or glycine. Northern blots  
 CC showed highest chitinase gene expression in lung and ovary tissues.  
 CC Expression in lung is consistent with a protective role against  
 CC pathogenic organisms that contain chitin. The invention provides  
 CC human chitinase polynucleotides and polypeptides, and materials and  
 CC methods for the recombinant production of human chitinase products,  
 CC which are expected to be useful as products for treating fungal  
 CC infections or for the development of such products. Human  
 CC chitinase has a synergistic effect on the actions of other  
 CC fungicides. It can be administered to improve the antifungal  
 CC activity of a non-chitinase antifungal agent, especially  
 CC amphotericin B or itraconazole, in the treatment of a fungal  
 CC infection such as candidiasis, aspergillosis, coccidioidomycosis,  
 CC blastomycosis, paracoccidioidomycosis, histoplasmosis,  
 CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,  
 CC dermatophytoses and Pneumocystis infections (all claimed). In  
 CC particular, the fungal infection involves Candida, Aspergillus  
 CC and/or Cryptococcus spp., whose growth is not effectively  
 CC inhibited by contact with human chitinase alone.

SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

## Alignment Scores:

Pred. No.: 2,886-200 Length: 1656  
 Score: 2392.00 Matches: 444  
 Percent Similarity: 99.78% Conservative: 0  
 Best Local Similarity: 99.78% Mismatches: 1  
 Query Match: 99.75% Indels: 0  
 DB: 24 Gaps: 0

US-10-004-219B-10 (1-445) x ABL57381 (1-1656)

Qy 1 AlalysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
 Db 90 GCAAAACTGGTCTGCTACTTACCACCACTGGGCCAGTAGTACAGACAGGGGAGGCTCGCTTC 149  
 Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40  
 Db 150 CTGCCCAAGGACTTTGGACCCCGCCAGCCCTTTGCACCCACCTCATCTACGCCCTTCCTGGCATG 209  
 Qy 41 ThrAsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 Db 210 ACCAACCCACGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACGAGAGTTCAAT 269  
 Qy 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
 Db 270 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTTAGCCATCGGAGCTGGAATTC 329  
 Qy 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
 Db 330 AGCACTCAGAAATTCACAGATATGTTAGCCAGGCCCAACACCCGTCAGACCTTTGTCAAC 389  
 Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
 Db 390 TGGGCCATCAGGTTTCTGGGCAATACAGCTTTGAGGCTTGACCTTGACTGGGAGTAC 449  
 Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrLeuValGlnAsp 140  
 Db 450 CCAGGAAGCCAGGGGAGCCCTGCGCTAGACAAGGAGCGCTTCACAAACCCCTGGTACAGGAC 509  
 Qy 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
 Db 510 TTGGCCAATGCTTCCAGCAGGAGGCCAGACCTCAGGAGAGGAACGCCCTTCTTCTGAGT 569  
 Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
 Db 570 GCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATAGAGGTGACAAATCGCC 629  
 Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200

```
Db 630 CAGAACCTGGATTTCACCTTATGGCTACGACTTCATGGCTCTTGGGAGAAAGGTC 689
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220
Db 690 ACGGACATAACAGCCCTCTACAGAGGCAAGAGAGTGGTCAGACGCCAGCTTC 749
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
Db 750 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAAGGGGACCCCTGCCAGCAAGCTGATC 809
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db 810 CTGGCATGCTACCTACCGAGCGCTCTTACACTGGCGCTCTCTCATCAGACACCAAGTG 869
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
Db 870 GGGGCCCCAGCCACAGGGTCTGGCATCTCAGGCCCTTCACCAAGAGAGGGGATGCTG 929
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 930 GCCTACTATGAAGTCTGCTCTGGAAGGGGGCCACCACACAGAGAATCCAGGATCAGAA 989
QY 301 ValProTyrIlePheArgAspGlnTrpValGlyPheAspValGluSerPheLys 320
Db 990 GTGCCCTTACATCTCCGGGACACCAAGTGGGTGGCTTTGATGATGGAGAGCTTCAAA 1049
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyValMetValTrpAlaLeuAsp 340
Db 1050 ACCNAGGTAGCTATCTGAAGCAGAGGAGACTGGGGGGGGCCCATGCTGGGCACTGGAC 1109
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCAGAGCTTA 1169
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1170 CGCGAGAACTGAGTCTTCATCTTGCCTTTCAGGCAACCCCGAGCTTGAAGTTCCAAA 1229
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1230 CAGGTTCAGCCCTCTGACCTGAGCATGGCCCGCCGAGCCCTGGACAGACAGCTTCTGCCAG 1289
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
Db 1290 GGCAGAGCTGATGGGCTCTATCCCAATCTCGGAAACGGTCCAGCTTCTACAGCTGTGCA 1349
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerIleSerCysLys 440
Db 1350 GCGGGCGGCTGTTCAGCAAAAGCTGCCGACAGGCGCTGGTGTTCAGCAACTCCTGCAAA 1409
QY 441 CysCysThrTrpAsn 445
Db 1410 TGCTGCACCTGGAT 1424
RESULT 10
AAT89181
ID AAT89181 standard; cDNA; 1768 BP.
XX
AC AAT89181;
XX
XX CysCysThrTrpAsn 445
XX
XX 27-APR-1998 (first entry)
XX
XX Human chitotriosidase variant cDNA.
XX
XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
XX
XX rheumatoid arthritis; atherosclerosis; human; ss.
XX
XX Homo sapiens.
XX
XX OS
XX Key Location/Qualifiers
XX CDS 124..1524
XX /*tag= a
XX
XX
```

```
PN W09736917-A1.
XX
XX 09-OCT-1997.
XX
XX 21-MAR-1997; 97WO-US05072.
XX
XX 29-MAR-1996; 96US-0014295.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;
XX P-PSDB; AAW31498.
XX
XX WPI; 1997-503041/46.
XX
XX New isolated human chitotriosidase gene - used to develop products
XX for the diagnosis and treatment of tissue remodeling disorders, e.g.
XX rheumatoid arthritis
XX
XX Claim 1; Page 21-22; 34pp; English.
XX
XX This human cDNA sequence encodes chitotriosidase (see AAW31498). It
XX is a preferred variant of the chitotriosidase cDNA sequence given
XX in AAT89180. Also claimed are: (1) a method of diagnosing a tissue
XX remodelling disorder related to expression of a mutated
XX chitotriosidase protein in a host comprising carrying out nucleic
XX acid amplification; and (2) a method of detecting altered
XX expression of a chitotriosidase protein in a host comprising
XX contacting a bodily sample with an antibody. Chitotriosidase can
XX degrade extracellular matrix substrates with a similar carbohydrate
XX structure to chitin. It can be used to develop products which can
XX be used in the diagnosis and treatment of tissue remodelling
XX disorders such as rheumatoid arthritis or atherosclerosis.
XX Chitotriosidase nucleic acids can be used in the recombinant
XX production of the enzyme, and as probes or primers in diagnostic
XX assays.
XX
XX Sequence 1768 BP; 401 A; 529 C; 474 G; 363 T; 1 other;
XX
Alignment Scores:
Pred. No.: 3,84e-200 Length: 1768
Score: 2391.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.71% Indels: 0
DB: 18 Gaps: 0
US-10-004-219B-10 (1-445) x AAT89181 (1-1768)
QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
Db 187 GCNAAACTGGTCTGCTACTTCCACCACTGGGCCCACTACAGACAGGGGAGCTCGCTTC 246
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 247 CTGCCCAAGGACTTGACCCAGCCCTTTGCCACCCACCTCATCTACGCTTCGCTGGCATG 306
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
Db 307 ACCAACCCAGCTGAGCACCACCTGAGTGGAGTACGAGACTCTCTACAGAGATTCAAT 366
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 367 GGCCTGAGAGAGATGAATCCCAAGCTGAAGCCCTGTTAGCCATCGAGGCTGGATTTC 426
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db 427 RGCACTCAGAAAGTTCCACAGATATGTCACCGGCAACCAACCGTCAGACCTTTGTCAAC 486
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 487 TCGGCCATCAGGTTTCTCGGCAATACAGCTTTGACGGCCTTGACCTTGAGTGGAGTAC 546
```

QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
DB 547 CAGGAAGCCAGGGAGCCCTGCGGTAGACAGGAGCGCTTCACAAACCCTGGTACAGAC 606  
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
DB 607 TTGGCCAAATGCTTCAGCAGGAAGCCAGACCTCAGGAGGAACGCGCTTCTTCAGT 666  
QY 161 AlaAlaValProAlaGlyGlnThrThrValAspAlaGlyTyrGluValAspLysLeuAla 180  
DB 667 GCAGCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACGAGTGGACAAATCGCC 726  
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTyrGluLysVal 200  
DB 727 CAGAACCTGGATTGTTCACCTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTC 786  
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220  
DB 787 ACGGACATACAGCCCTCTACAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCCTC 846  
QY 221 AsnValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuLeu 240  
DB 847 AACGTGGATGCTGTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATC 906  
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
DB 907 CTTGGCATGCTTACCTACGAGCGCTCTTCACCTACCTGGCTCTCATCAGACACAGAGTG 966  
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
DB 967 GGGGCCCCAGCCACAGGGTCTGGCATCTCCAGGCCCTTCACCAAGCAAGAGGATGCTG 1026  
QY 281 AlaTyrTyrGluValCysSerTyrLysGlyAlaThrLysGlnArgLleGlnAspGlnLys 300  
DB 1027 GCCTACTATGAAGTCTGCTCTCGAAGGGGGCCACCAACAGAGAATCCAGATCAGAG 1086  
QY 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspValGluSerPheLys 320  
DB 1087 GTGCCCTAGCTCTTCGGGACCAACCGTGGGTGGCTTGTATGATGTGGAGAGCTTCAAA 1146  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLysGlyAlaMetValTrpAlaLeuAsp 340  
DB 1147 ACCAAGGTCAAGTATCTGAAGCAGAGGAGTGGCGGGGCCATGCTCTGGGCACTGGAC 1206  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
DB 1207 TTAGATGACTTTGGCGGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCCAGACGCTA 1266  
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuValProLys 380  
DB 1267 CGGCAGGAAGTGTCTTCATCTTCTTCCAGCACCCAGAGCTTGAAGTTCAGAAA 1326  
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
DB 1327 CCAGGTCAAGCTCTGAACCTGAGCATGCCGCCAGCCCTGGACAGACACGCTTCTGCCAG 1386  
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
DB 1387 GGCAGAGCTGATGGGCTCTATCCCAATCTCCGGAACGGTCCAGCTTCTACAGCTGTGCA 1446  
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
DB 1447 GCGGGGCGCTGTTCAGCAAGAGTGGCCGACAGAGCGCTGTGTTCAGCAACTCTTCGAAA 1506  
QY 441 CysCysThrTrpAsn 445  
DB 1507 TGCTGCACCTGGAAT 1521

RESULT 11

AAT89180

ID AAT89180 standard; cDNA; 1637 BP.

XX

AC AAT89180;  
XX 27-APR-1998 (first entry)  
XX Human chitotriosidase cDNA.  
XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;  
KW rheumatoid arthritis; atherosclerosis; human; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH CDS 27..1421  
FT /\*tag= a  
FT variation 330  
FT /\*tag= b  
FT /note= "nucleotide 330 is G, or an insertion of  
FT GGTACA, in some variants of the sequence"  
FT variation 1253  
FT /note= "nucleotide 1253 is C in some variants of  
FT the sequence"  
FT variation 1426  
FT /note= "nucleotide 1426 is deleted in some variants  
FT of the sequence"  
FT variation 1442  
FT /note= "nucleotide 1442 is deleted in some variants  
FT of the sequence"  
FT variation 1495  
FT /note= "nucleotide 1495 is deleted in some variants  
FT of the sequence"  
FT variation 1632  
FT /note= "nucleotide 1632 is G in some variants of  
FT the sequence"  
FT variation 1634  
FT /note= "nucleotide 1634 is G in some variants of  
FT the sequence"  
XX WO9736917-A1.  
XX 09-OCT-1997.  
XX 21-MAR-1997; 97WO-US05072.  
XX 29-MAR-1996; 96US-0014295.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (SMIK) SMITHKLINE BEECHAM CORP.  
XX Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;  
XX WPI: 1997-503041/46.  
XX P-PSDB; AAW31497.  
XX New isolated human chitotriosidase gene - used to develop products  
XX for the diagnosis and treatment of tissue remodeling disorders, e.g.  
XX rheumatoid arthritis  
XX Claim 1; Page 18-19; 34pp; English.  
XX This human cDNA sequence includes a coding region of  
XX chitotriosidase (see AAW31497). It is deposited as ATCC 69953.  
XX Variants of the sequence are provided, including the preferred  
XX sequence variant given in AAT89180. Also claimed are: (1) a method  
XX of diagnosing a tissue remodelling disorder related to expression  
XX of a mutated chitotriosidase protein in a host comprising carrying  
XX out nucleic acid amplification; and (2) a method of detecting  
XX altered expression of a chitotriosidase protein in a host

CC comprising contacting a bodily sample with an antibody.  
 CC Chitinase can degrade extracellular matrix substrates with a  
 CC similar carbohydrate structure to chitin. It can be used to  
 CC develop products which can be used in the diagnosis and treatment  
 CC of tissue remodelling disorders such as rheumatoid arthritis or  
 CC atherosclerosis. Chitinase can degrade nucleic acids can be used in the  
 CC recombinant production of the enzyme, and as probes or primers in  
 CC diagnostic assays.

XX  
 SQ Sequence 1637 BP; 355 A; 493 C; 443 G; 346 T; 0 other;

## Alignment Scores:

Pred. No.: 1.61e-198 Length: 1637  
 Score: 2372.00 Matches: 442  
 Percent Similarity: 99.33% Conservative: 0  
 Best Local Similarity: 99.33% Mismatches: 1  
 Query Match: 98.92% Indels: 2  
 DB: 18 Gaps: 1

US-10-004-219b-10 (1-445) x AAT89180 (1-1637)

QY 1 AlaLysLeuValCysTyrPheThrAsnTTPAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
 DB 90 GCAAAACTGGTCTGCTACTTTCACCAACTGGGCCAGTACAGAGGGGAGGCTCGCTTC 149  
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
 DB 150 CTGCCCAAGGACTTGGACCCCGCCCTTTCACCCACCTCATCTACGCCCTTCGCTGCATG 209  
 QY 41 ThrAsnHisGlnLeuSerThrThrGluTTPAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 DB 210 ACCAACCCAGCTGAGCACCCTGAGTGGGAATGAGCAGACTCTTACCAGGAGTTCAAT 269  
 QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTTPAsnPhe 80  
 DB 270 GGCTGAAGAAGATGAATCCCAAGCTGAGAGCCCTGTAGCCATCGAGGCTGGAATTC 329  
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
 DB 330 AGCACTCAGAAGTTACAGATATGGTAGCCAGCCGCAACACCGTCAGACCTTTGTCAAC 389  
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTTPGluTyr 120  
 DB 390 TCGGCCATCAGGTTCTTGGCCAAATACAGCTTTGACGGCCCTTGACCTTGACCTGGAGTAC 449  
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
 DB 450 CAGGAAGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCACACCCCTG-----GAC 503  
 QY 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160  
 DB 504 TTGGCCAAATGCCCTTCAGCAGGAAGCCAGACCTCAGGAAGGAACGCCCTTCTCTGAGT 563  
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
 DB 564 GCAGCGGTTCCAGCTGGCAGACCTATGTGGATGCTGATACAGAGTGGAGCAAAAATCGCC 623  
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTTPGluLysVal 200  
 DB 624 CAGAACCTGGATTTGTCAACCTTATGGCTCAGACTTCATGGCTCTTGGGAAGAGGTC 583  
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyAlaAlaSerLeu 220  
 DB 684 ACGGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGCAGCAGCCCTC 743  
 QY 221 AsnValAspAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIle 240  
 DB 744 AACGTGGATGCTCTGTGCAACAGTGGCTGCAAGGGGACCCCTCCAGCAGCTGATC 803  
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
 DB 804 CTTGGCATGCTTACCTACGGAACCTCTCTTCACACTGGCTCTCTCANTACAGACAGAGTG 863

QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
 DB 864 GGGGCCCCCAGCCACAGGGTCTGGCACTCCAGGCCCTTCCACCAAGAAAGGAGGATGCTG 923  
 QY 281 AlaTyrTyrGluValCysSerTTPAsnGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
 DB 924 GCCTACTATGAAGTCTGCTCTGGAAAGGGGGCCACCAACAGAGAATCCAGATCAGAG 983  
 QY 301 ValProTyrIlePheArgAspAsnGlnTTPValGlyPheAspAspValGluSerPheLys 320  
 DB 984 GTGCCCTACATCTTCGGGACCAACAGTGGGTGGGCTTTGATGATGGAGAGCTTCAA 1043  
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTTPAlaLeuAsp 340  
 DB 1044 ACCAAGGTCAGCTATCTGAAGCAGAAGGAGCTGGGGGGGGCCATGGTCTGGGCACTGGAC 1103  
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
 DB 1104 TTAGATGACTTTTGGCGGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCCAGAGCTA 1163  
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
 DB 1164 CGGCAGGAACTGAGTCTTCCATCTTGGCTTCAGGCACCCAGAGCTTGAAGTTCACAAA 1223  
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
 DB 1224 CCAGGTCAGCCCTCTGAACCTGAGCATGGTCCAGCCCTGGCAAGACACGCTTCTGCCAG 1283  
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
 DB 1284 GGCAAGCTGATGGGCTCTATCCCAATCTCGGAACCGGTCCAGCTTCTACAGCTGTGCA 1343  
 QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
 DB 1344 GGGGGGGGCTGTTCCAGCAAGCTGCCGACAGGCTGGTTCAGCACTCTCTGCAACTCTT 1403  
 QY 441 CysCysThrTTPAsn 445  
 DB 1404 TGCTGCACCTGGAAT 1418

RESULT 12  
 AAT50834  
 ID AAT50834 standard; cDNA; 1713 BP.  
 XX  
 AC AAT50834;  
 XX  
 DT 24-MAR-1997 (first entry)  
 XX  
 DE Human chitinase cDNA clone chi.39.  
 XX  
 KW Chitinase; chitinase; chitin; infectious disease;  
 KW gene therapy; vaccine; lysosomal lipodosis; Gaucher's disease;  
 KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;  
 KW multiple sclerosis; drug delivery; cosmetics; food; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 13..1176  
 FT /tag= a  
 FT sig\_peptide 13..75  
 FT /tag= b  
 FT mat\_peptide 76..1173  
 FT /tag= c  
 XX  
 PN WO9640940-A2.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 06-JUN-1996; 96WO-NL00225.  
 XX  
 PR 07-JUN-1995; 95US-0486839.  
 XX

(UNAM ) UNIV AMSTERDAM.

Aerts JWFg;

WPI: 1997-118698/11.

P-PSDB; AAW08585.

New human chitinase - used to treat or prevent infection by  
chitin-contg. pathogens, in diagnosis and as additives to cosmetics,  
foods, implants etc.

Claim 2; Page 42-43; 59pp; English.

A cDNA clone (AA050834), designated chi.39, codes for a 39 kDa human  
chitinase (AAW08584) that is stable to many proteases, active at pH 3-  
8 and up to 50 deg, and stable in the circulation. Clones chi.39  
and chi.50 (see also AA050833) were isolated from a human macrophage  
cDNA library using as probe a partial clone obtd. using primers  
(see also AA050835-36) based on a chitotriosidase purified from a  
type 1 Gaucher disease patient. The 2 cDNA clones are the result  
of alternative splicing of RNA. Chitinase nucleic acid can be used  
for large-scale prodn. of recombinant human chitinases, or can be  
incorporated into a gene therapy vector to treat or prevent  
infection by chitin-contg. pathogens.

Sequence 1713 BP; 380 A; 503 C; 465 G; 365 T; 0 other;

## Alignment Scores:

Pred. No.:	1.98e-196	Length:	1713
Score:	2348.50	Matches:	443
Percent Similarity:	94.66%	Conservative:	0
Best Local Similarity:	94.66%	Mismatches:	1
Query Match:	97.94%	Indels:	25
DB:	18	Gaps:	1

US-10-004-219B-10 (1-445) x AAT50834 (1-1713)

QY	2	LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu	21
DB	79	AAACTGGTCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTCCTG	138
QY	22	ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr	41
DB	139	CCCAAGGACTTGGACCCAGCCCTTTGACCCACCTCATCTAGCCCTTCGCTGGCATGCC	198
QY	42	AsnHisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly	61
DB	199	AACCCACCTGAGCACCCTGAGTGAATGACGAGACTCTTACCAGAGGTCAATGGC	258
QY	62	LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly	81
DB	259	CTGAAGAGATGATATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAAATTCGCG	318
QY	82	ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer	101
DB	319	ACTCAAGATTTCACGATATGTTAGTACCGGCCAACACCGTCAGACCTTTTGTCAACTCG	378
QY	102	AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro	121
DB	379	GCCATCAGGTTTCTGGCAATATACAGCTTTGACGGCTTGTAGCTTGGAGTACCCA	438
QY	122	GlySerGlnGlySerProAlaValaAspLysGluArgPheThrThrLeuValGlnAspLeu	141
DB	439	GGAGCCAGGGAGCCCTCGGTAGACAGAGAGCGCTTCACAACTCGGTGACAGACTTG	498
QY	142	AlaAsnAlaPheGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla	161
DB	499	GCCAAATGCTTCAGCAGGAGGCCAGACTCAGGAAGAAACGCTTCTTCTGAGTGCA	558
QY	162	AlaValProAlaGlyGlnThrThrValAspAlaGlyTyrGluValAspLysIleAlaGln	181
DB	559	GCGGTTCAGCTGGGCAGACCTATGTGGATGTGGATACGAGGTGACAAATGCCCCAG	618

QY	182	AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr	201
DB	619	AACCTGGATTTTGTCAACCTTATGSCCTTACGACTTCCATGCTCTTGGAGAGTCAAG	678
QY	202	GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAsnSerLeuAsn	221
DB	679	GGACATAACAGCCCTCTACAGAGGCAAGAAGAGTGTGTGCAGCAGCCAGCTCAAC	738
QY	222	ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu	241
DB	739	GTGATGCTGTGTGCAACAGTGTGCAAGAGGGGACCTCTGCCAGCAGCTGATCTT	798
QY	242	GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly	261
DB	799	GGCATGCTTACTACGGGAGCTCTTTCACACTGGGCTCTCATCAGACACAGAGTGGG	858
QY	262	AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla	281
DB	859	GCCCCAGCCACAGGGTCTGGGCACTCCAGGCCCTTCCACCAAGGAAGAGGGATGCTGCC	918
QY	282	TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal	301
DB	919	TACTATGAAGTCTGCTCTCGAAGGGGGCCACCAACAGAGNATCCAGATCAGAGGTG	978
QY	302	ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThr	321
DB	979	CCCTACATCTTCCGGGACCAACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAAAACC	1038
QY	322	LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu	341
DB	1039	AAGGTCAGCTATCTGAAGCAGAGGGAGTGGGGCGGGCCATGGTCTGGGCACCTGGACTTA	1098
QY	342	AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg	361
DB	1099	GATGACTTTGGCGGCTTCTCTGCAACCCAGGGCCGATACCCCTCATCCAGACGCTACGG	1158
QY	362	GlnGluLeuSer-----	365
DB	1159	CAGGAACCTGAA- TGGGTAAAGCCTTAACTGCTGTGCATGTGAGGCCAGGTTTGCCTG	1217
QY	366	-----LeuProTyrLeuProSerGlyThrProGluLeuGlu	377
DB	1218	TGGCACTGTGCTTTCAGCTGTAGTCTTCCATCTTTCAGGCACCCAGAGTTGAA	1277
QY	378	ValProLysProGlyGlnProSerGlnProGluHisGlyProSerProGlyGlnAspThr	397
DB	1278	GTTCCAAAACAGGTCAGCCCTCTGAACTGAGCATGGCCCAAGCCCTGGACACACAGC	1337
QY	398	PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyr	417
DB	1338	TTCTGCCAGGGCAAGCTGTATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTAC	1397
QY	418	SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn	437
DB	1398	AGCTGTGAGGGGGGGCTGTTCAGCAAAAGTGTCCCGCAGACGGCTGTGTTCAGCAAC	1457
QY	438	SerCysLysCysCysThrTrpAsn 445	
DB	1458	TCCTGCAANTGTGCACCTGGAAT 1481	
RESULT 13			
AAH42013			
ID	AAH42013 standard; DNA; 1368 BP.		
XX	AAH42013;		
AC	AAH42013;		
XX	24-AUG-2001 (first entry)		
DT	Disease treatment related oligonucleotide SEQ ID NO: 3.		
DE	Disease treatment; infection; chronic occlusive pulmonary disease;		
KW	bronchial asthma; ds.		
XX			

OS Homo sapiens.  
 XX WO200136633-A1.  
 XX PD 25-MAY-2001.  
 XX PF 14-NOV-2000; 2000WO-JP08015.  
 XX PR 15-NOV-1999; 99JP-0324467.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX Nakanishi A, Morita S;  
 XX WPI; 2001-397791/42.  
 XX New proteins, peptides and DNA for treatment of bronchial asthma,  
 PT chronic occlusive lung disease and infectious disease  
 XX Claim 5; Page 100; 114pp; Japanese.  
 CC The present invention provides the sequence of a protein which can be  
 CC used in the treatment and prevention of infectious diseases. Inhibitors  
 CC of the protein can be used to treat bronchial asthma and chronic  
 CC occlusive pulmonary disease. The present sequence is an oligonucleotide  
 CC described in the exemplification of the invention.  
 XX SQ Sequence 1368 BP; 324 A; 398 C; 349 G; 297 T; 0 other;

Alignment Scores:  
 Pred. No.: 1,42e-106 Length: 1368  
 Score: 1323.00 Matches: 240  
 Percent Similarity: 67.39% Conservative: 70  
 Best Local Similarity: 52.17% Mismatches: 126  
 Query Match: 55.17% Indels: 24  
 DB: 22 Gaps: 3

US-10-004-219B-10 (1-445) x AAH42013 (1-1368)

QY 2 LysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
 DB 4 CAGCTGACATGCTACTTCAACAACTGGGCCAGTACCGGCCAGGCTGGGGCCCTTCATG 63  
 QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
 DB 64 CCTGACACATCGACCCCTCTGTACCCACCTGATCTACGCTTGTGCGAGGCGAG 123  
 QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
 DB 124 AACACAGAGATCACCAACCATCGAATGGAATGATGATGCTCTTACCAAGCTTTCAATGGC 183  
 QY 62 LeuLysIleMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrPheAsnGly 81  
 DB 184 CTGAAATAAAGAACACAGCCAGCTGAAACTCTCTGGCCATTGGAGGCTGGAACTTCGGG 243  
 QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
 DB 244 ACTGCCCTTTTCACTGCCATGGTTTCTACTCTCTGAGACCGCCAGACTTTCATCACCCTCA 303  
 QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspThrGluTyrPro 121  
 DB 304 GTCATCAAAATCTCGGCCAGTATGATGTTGACGGGCTGGACTTTTCACTGGGAGTACCCT 363  
 QY 122 GlySerGlnCysSerProAlaValAspLysGluArgPheThrThrThrLeuValGlnAspLeu 141  
 DB 364 GGCTCTCGGGAGCCCTCTCTAGGACAGCATCTTCACTGCTCTGGTGCGAGAAATG 423  
 QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
 DB 424 CGTGAAGCTTTTTCAGCAGGAGGAGCAAGCAGATCAACAGCCAGGCTGATGGTCACTGCT 483  
 QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181

DB 484 GCAGTAGCTGCTGGCATCTCCAATATCCAGTCTGCTATGAGATCCCCCAACTGTCACAG 543  
 QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTyrGluLysValThr 201  
 DB 544 TACCTGGACTACATCATGTCATGACCTACGACCTCCATGGCTCTGGAGGGCTACACT 603  
 QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeuAsn 221  
 DB 604 GGAGAGAAACAGCCCTCTCAAAATACCGACTGACACCGGACGCAAGCCCTACCTCAAT 663  
 QY 222 ValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuLeu 241  
 DB 664 GTGGATTATGTCATCACTACTGGAAGGACCAATGACGACCACTGAGAGCTCATCGT 723  
 QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261  
 DB 724 GGATTCCTTACCTATGGACACAACTTCTCTGACCAACCCCTCCACACATGGAATTGT 783  
 QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281  
 DB 784 GCCCCCACCTCTGGTGTCTGCTGCGCCCTATGCCAAGGAGCTCTGGATCTGGGCT 843  
 QY 282 TyrTyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGln 299  
 DB 844 TACTACGAGATCTGTACCTTCTGAAATGAGGACCACTCAGGATGGATGCCCTCAG 903  
 QY 300 LysValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPhe 319  
 DB 904 GAAGTGCTTATGCTATCAGGGCAATGTGGTGGTGGCTATGACCAACATCAAGAGCTTC 963  
 QY 320 LysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeu 339  
 DB 964 GATATTAAAGGCTCAATGGCTTAAGCACAAATTTGGAGGCGCCATGCTCTGGGCCATT 1023  
 QY 340 AspLeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThr 359  
 DB 1024 GATCTGGATGACTTCTACCTGGCATTCTTGCACACAGGGCAAGTTTCCCCCTAATCTCCACC 1083  
 QY 360 LeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro 379  
 DB 1084 CTGAGAGAGGCCCTC-----GGCTGCAGAGTGAAGTTGCACG 1122  
 QY 380 LysProGlyGlnProSerGluPro----- 387  
 DB 1123 GCTCCAGCTCAGCCCATTCAGCCAAATACTGCTGCTCCAGTGGCAGCGGGAACGGAGC 1182  
 QY 388 -----GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAsp 404  
 DB 1183 GGGAGTAGCAGCTCTGGAGGCGAGCTCGGAGGCGAGTGGATTCTGTGCTGCAGAGCCAAAC 1242  
 QY 405 GlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAlaAlaGlyArgLeu 424  
 DB 1243 GGCCTCTACCCCGTGGCAATAACAGAAATGCTTCTGGCACTGGCGTGAATGAGTACAG 1302  
 QY 425 PheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCysThrTrp 444  
 DB 1303 TACCAGCAGAACTGCAGCGCGGCTGTCTTCGACACAGCAGCTGTGATTCGTGCAACTGG 1362

RESULT 14  
 AAH42025  
 ID AAH42025 standard; DNA; 1678 BP.  
 XX AC AAH42025;  
 XX DT 24-AUG-2001 (first entry)  
 XX DE Disease treatment related oligonucleotide SEQ ID NO: 16.  
 XX KW Disease treatment; infection; chronic occlusive pulmonary disease;  
 XX OS bronchial asthma; ds.  
 XX OS Homo sapiens.



PN WO200136633-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-JP08015.

XX 15-NOV-1999; 99JP-0324467.

XX (TAKE ) TAKEDA CHEM IND LTD.

PA Nakanishi A, Morita S;

XX WPI; 2001-397791/42.

XX New proteins, peptides and DNA for treatment of bronchial asthma,

XX chronic occlusive lung disease and infectious disease

PI Example 5; Page 107-108; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be

XX used in the treatment and prevention of infectious diseases. Inhibitors

XX of the protein can be used to treat bronchial asthma and chronic

XX occlusive pulmonary disease. The present sequence is an oligonucleotide

XX described in the exemplification of the invention.

XX Sequence 1678 BP; 397 A; 497 C; 418 G; 366 T; 0 other;

SQ

Alignment Scores:

Pred. No.: 1,86e-106 Length: 1678

Score: 1323.00 Matches: 240

Percent Similarity: 67.39% Conservative: 70

Best Local Similarity: 52.17% Mismatches: 126

Query Match: 55.17% Indels: 24

DB: 22 Gaps: 3

US-10-004-219B-10 (1-445) x AAH42025 (1-1678)

QY 2 LysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21

DB 209 CAGTGACATGCTACTTCAACCACTGGGCCCACTACCGCCAGGCTGGGGCGCTTCATG 268

QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41

DB 269 CCGTCAACATCGACCCCTGCTGTATCCACCTGATCCACCTGATCCGCTTTCATGAGGCG 328

QY 42 AsnHisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnLysPheAsnGly 61

DB 329 AACACGAGATCACCACCACTCAAGTGAATGATGTGACTCTTACCAAGCTTTCATATGGC 388

QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81

DB 389 CTGAAAATAAGACACAGCCAGCTGAACACTCTCTGGCCATTGGAGGCTGGAACTTCGGG 448

QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101

DB 449 ACTGCCCTTTTCACTGCCATGGTTTCTACTCTCGAGAACCGCCAGACTTTCATCACCTCA 508

QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121

DB 509 GTTCATCAAAATCTCGGCCAGTATGATGTTGACGGGCTGGACTTTGACTGGGAGTACCT 568

QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141

DB 569 GGCTCTCGTGGAGCCCTCTCAGACAGCATCTTCACTGTCTGTGGTGGAGAAATG 628

QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161

DB 629 CGTGAAGCTTTTGAAGAGGAGGCAAGAGATCAACAGCCAGGCTGATGTCTACTGTCT 688

QY 162 AlaValProAlaGlyGlnThrThrValAspAlaGlyTyrGluValAspLysIleAlaGln 181

DB 689 GCAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTACAG 748

QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201

DB 749 TACCTGGACTACATCCATGCTACGACCTTCCATGCTCTCTGGAGGCTACACT 808

QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyAlaAlaAsnSerLeuAsn 221

DB 809 GGAGAGAAACAGCCCTCTACAAATACCCGACTGACACCGCAGCAACGCCCTTACCTCAAT 868

QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241

DB 869 GTGGATTATGTCATGAACACTCTGGAGGACAAATGGACCACTGAGAGCTCATCGTT 928

QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261

DB 929 GGAATTCCTTACCTATGGACACAACTTTCATCTGAGCAACCCCTCCACACTGGAATTTGT 988

QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281

DB 989 GCGCCACCTCTGGTGTCTGCTGGGCCCTATGCCAAGGAGTCTGGGATCTGGGCT 1048

QY 282 TyrTyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGln 299

DB 1049 TACTACGAGATCTGTACCTTCTGAAAAATGAGGCACTCAGGATGGATGCCCTCAG 1108

QY 300 LysValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPhe 319

DB 1109 GAAGTGCCTTATGCTATCAGGCAATGTGTGGTGTGCTATGACACATCAAGAGCTTC 1168

QY 320 LysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeu 339

DB 1169 GATATTAAAGGCTCAATGGCTTAAGCACAAATAATTGGAGGCGCCATGCTCTGGGCCATT 1228

QY 340 AspLeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThr 359

DB 1229 GATCTGATGATCTTCACTGGCATTCTTGTCAACACGAGGCAAGTTTCCCTAATCTCCACC 1288

QY 360 LeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro 379

DB 1289 CTGAAGAAGGCCCTC-----GGCTCGACAGTGCAAGTTGCACG 1327

QY 380 LysProGlyGlnProSerGluPro----- 387

DB 1328 GCTCAGCTCAGCCCATTTGAGCCAAATAACTGCTGCCAGTGGCAGCGGGAACCGGAGC 1387

QY 388 -----GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAsp 404

DB 1388 GGGAGTAGCAGCTCTGGAGGAGGAGCTCGGAGGAGGAGTGTCTGTCTGAGAGCCAAAC 1447

QY 405 GlyLeuTyrProAsnProArgGluArgSerPheTyrSerCysAlaAlaGlyArgLeu 424

DB 1448 GGCCTCTACCCCGTGGCAATAACAGAAATGCCCTTCTGGCACTGCGTGAATGGAGTCACG 1507

QY 425 PheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCysThrTrp 444

DB 1508 TACCAGCAGAACTGCCAGCGCGGCTTCTCTCGACACCACTGATGTGCTGCAACTGG 1567

RESULT 15

ABNS9993

ID ABNS9993 standard; cDNA; 1369 BP.

XX AC ABNS9993;

XX DT 28-JUN-2002 (first entry)

XX DE Novel human coding sequence SEQ ID NO: 404.

XX KW Human; antianemic; vulnerary; antiinflammatory; immunomodulator;

XX KW antifertility; cerebrotective; cytostatic; rheumatic; gene therapy;

XX KW neuroprotective; antiparkinsonian; protein therapy; EST;

XX XX expressed sequence tag; gene; ss.

XX OS Homo sapiens.

PN W0200222660-A2.  
XX 21-MAR-2002.  
XX 10-SEP-2001; 2001WO-US26015.  
XX 11-SEP-2000; 2000US-0659671.  
XX (HYSE-) HYSEQ INC.  
XX PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
XX PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX DR WPI; 2002-292408/33.  
XX DR P-PSDB; ABB97580.  
XX PT An isolated polynucleotide for treating diseases associated with its  
XX encoded polypeptide such as cancer and multiple sclerosis -  
XX Claim 1; SEQ ID NO 404; 509pp; English.  
XX CC The present invention provides the protein and coding sequences of 444  
XX novel human proteins. These were isolated from expressed sequences tags  
XX (ESTs). They can be used to stimulate cell growth, to regulate  
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
XX e.g. in burn treatment, to regulate the immune system e.g. to treat  
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat  
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat  
XX stroke and cancer, to screen for drugs, to treat inflammatory conditions  
XX e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
XX Parkinson's disease. The present sequence is a coding sequence of the  
XX invention.  
XX SQ Sequence 1369 BP; 327 A; 393 C; 341 G; 308 T; 0 other;

Alignment Scores:  
Pred. No.: 1,36e-84 Length: 1369  
Score: 1072.50 Matches: 200  
Percent Similarity: 64.53% Conservative: 62  
Best Local Similarity: 49.26% Mismatches: 117  
Query Match: 44.72% Indels: 27  
DB: 24 Gaps: 4

US-10-004-219B-10 (1-445) x ABB59993 (1-1369)

QY 56 TyrGlnGluPheAsnGlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIle 75  
DB 67 TATACTGAATTT-----GCAGCTCGGAACAGCCAGCTGAAAACTCTCTGGCCATT 117  
QY 76 GlyGlyTyrAsnPheGlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArg 95  
DB 118 GGAGGCTGGAACTTCGGGACTGCCCTTTCACCTGCCATGGTTCTACTCTCTGAGAACCGC 177  
QY 96 GlnThrPheValAsnSerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAsp 115  
DB 178 CAGACTTTCATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 237  
QY 116 LeuAspTyrGluTyrProGlySerGlnGlySerProAlaValAspLysGluArgPheThr 135  
DB 238 TTTGACTGGGAGTACCTGGCTCTCGTGGAGCCCTCCTCAGGACCAAGCATCTCTTCACT 297  
QY 136 ThrLeuValGlnAspLeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGlu 155  
DB 298 GTCTGTGTGCGGAATGCGTGAAGCTTTTGACGAGGAGCCCAAGCAGATCAACAGCCCC 357  
QY 156 ArgLeuLeuSerAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGlu 175  
DB 358 AGGCTGATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 417  
QY 176 ValAspLysIleAlaGlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGly 195  
DB 418 ATCCCCCAACTGTACAGTACTCTGGAGTACATCCATGTGATGACCTACGACCTCCATGGC 477

QY 196 SerTyrGluLysValThrGlyHisSerProLeuTyrLysArgGlnGluGluSerGly 215  
DB 478 TCCTGGAGGGCTACACTGGAGAGAACAGCCCTCTACAAATACCCGACTGACACCGGC 537  
QY 216 AlaAlaAlaSerLeuAsnValAspAlaValGlnGlnThrLeuGlnLysGlyThrPro 235  
DB 538 AGCAACGCTCACTCAATGTGATTTATGTCATGAATCTGGAAGGACAATGGAGCACCA 597  
QY 236 AlaSerLysLeuLeuLeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSer 255  
DB 598 GCTGAGAGCTCATCTGTTGGATTCCTACTATGACACAACTTCTCTGAGCAACCCC 657  
QY 256 SerAspThrArgValGlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLys 275  
DB 658 TCCAACTGGAATGGTGGCCCACTCTGGTGGTCTGCTGCTGGCCCTATGCCAAG 717  
QY 276 GluGlyGlyMetLeuAlaTyrGluValCysSerTrp-----LysGlyAlaThrLys 293  
DB 718 GAGTCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTCGTGAATAATGGAGCCACTCAG 777  
QY 294 GlnArgIleGlnAspGlnLysValProTyrIlePheArgAspAsnGlnTrpValGlyPhe 313  
DB 778 GGATGGATGCCCTCAGGAAGTGCCTTATGCTATCAGGCAATGTGTGGTTGGCTAT 837  
QY 314 AspAspValGluSerPheLysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGly 333  
DB 838 GACAACTCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAAATTTGGAGGC 897  
QY 334 AlaMetValTrpAlaLeuAspLeuAspPheAlaGlyPheSerCysAsnGlnGlyArg 353  
DB 898 GCCATGGTCTGGCCCAATGATCTGGATCACTTCTGCAACAGGCGCAAG 957  
QY 354 TyrProLeuIleGlnThrLeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThr 373  
DB 958 TTTCCCTTATCTCCACCTGAAGAGCCCTC-----GGCCTG 996  
QY 374 ProGluLeuGluValProLysProGlyGlnProSerGluPro----- 387  
DB 997 CAGAGTGCAAGTTGCACGGCTCCAGCTCAGCCCAATTGAGCCAAATACTGCTCCTCCACT 1056  
QY 388 -----GluHisGlyProSerProGlyGlnAspThrPhe 398  
DB 1057 GGCACGGGAACGGGAGCGGAGTAGCAGCTCTGAGGAGGAGCTCGGAGGAGGAGTTC 1116  
QY 399 CysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSer 418  
DB 1117 TGTGCTGTGACAGGCAACGGGCTCTACCCCGTGGCAATAACAGAAATGCCTTCTGGCAC 1176  
QY 419 CysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSer 438  
DB 1177 TGCGTGAATGGAGTCACGTACCAGCAGAACTGCCAGGCGGGCTTGTCTTCGACACCCAGC 1236  
QY 439 CysLysCysCysThrTrp 444  
DB 1237 TGTGATTCGTCAACTGG 1254

Search completed: June 29, 2003, 21:27:02  
Job time : 266.223 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: June 29, 2003, 21:12:38 ; Search time 63.5244 Seconds  
(without alignments)  
2148.326 Million cell updates/sec

Title: US-10-004-219B-10  
Perfect score: 2398  
Sequence: 1 AKLVCFYTNWAGYRQGEARF.....QQSCPTGLVFNCKCKCTWN 445

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlpl  
-Q=/cgn2\_1/USPTO.spool/US10004219/runat\_24062003\_160229\_8184/app\_query.fasta\_1.1877  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219@cgn\_1\_1\_108@runat\_24062003\_160229\_8184 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	100.0	1636	4	US-09-039-198A-1
2	2398	100.0	1636	4	US-08-877-599-1
3	2398	100.0	1636	4	US-09-267-574-1
4	2394	99.8	1643	2	US-08-486-839-3
5	2394	99.8	1643	3	US-09-151-011-3
6	2394	99.8	1643	4	US-09-343-623-3
7	2392	99.7	1656	4	US-09-039-198A-3
8	2392	99.7	1656	4	US-08-877-599-3
9	2392	99.7	1656	4	US-09-267-574-3
10	2348.5	97.9	1713	2	US-08-486-839-5
11	2348.5	97.9	1713	3	US-09-151-011-5
12	2348.5	97.9	1713	4	US-09-343-623-5

13	1046.5	43.6	1433	1	US-08-694-915-1	Sequence 1, Appl
14	1046.5	43.6	1526	1	US-08-694-915-3	Sequence 3, Appl
15	1030	43.0	1681	5	PCT-US94-07754-4	Sequence 4, Appl
16	1025	42.7	1681	2	US-08-581-527-4	Sequence 4, Appl
17	824	34.4	2452	2	US-08-524-051-1	Sequence 1, Appl
18	803.5	33.5	1677	4	US-09-545-814-13	Sequence 13, Appl
19	803.5	33.5	1677	4	US-09-545-814-15	Sequence 15, Appl
20	803.5	33.5	1749	4	US-09-545-814-4	Sequence 4, Appl
21	803.5	33.5	1749	4	US-09-545-814-6	Sequence 6, Appl
22	803.5	33.5	2610	4	US-09-545-814-1	Sequence 1, Appl
23	803.5	33.5	2610	4	US-09-545-814-3	Sequence 3, Appl
24	797.5	33.3	1919	4	US-09-545-814-31	Sequence 31, Appl
25	797.5	33.3	1919	4	US-09-545-814-33	Sequence 33, Appl
26	762	31.8	1470	4	US-09-292-225-40	Sequence 40, Appl
27	762	31.8	1470	4	US-09-292-225-42	Sequence 42, Appl
28	762	31.8	1527	4	US-09-292-225-37	Sequence 37, Appl
29	762	31.8	1527	4	US-09-292-225-39	Sequence 39, Appl
30	762	31.8	1621	4	US-09-292-225-34	Sequence 34, Appl
31	762	31.8	1621	4	US-09-292-225-36	Sequence 36, Appl
32	738	30.8	1608	4	US-09-292-225-20	Sequence 20, Appl
33	738	30.8	1608	4	US-09-292-225-22	Sequence 22, Appl
34	738	30.8	1665	4	US-09-292-225-17	Sequence 17, Appl
35	738	30.8	1665	4	US-09-292-225-19	Sequence 19, Appl
36	738	30.8	1752	4	US-09-292-225-14	Sequence 14, Appl
37	738	30.8	1752	4	US-09-292-225-16	Sequence 16, Appl
38	660	27.5	1478	4	US-09-545-814-28	Sequence 28, Appl
39	660	27.5	1478	4	US-09-545-814-30	Sequence 30, Appl
40	499.5	20.8	1167	1	US-07-939-501A-6	Sequence 6, Appl
41	499.5	20.8	1167	4	US-08-448-398-10	Sequence 10, Appl
42	499.5	20.8	1320	1	US-07-939-501A-14	Sequence 14, Appl
43	499.5	20.8	1364	1	US-07-939-501A-13	Sequence 13, Appl
44	499.5	20.8	1405	1	US-07-939-501A-9	Sequence 9, Appl
45	486	20.3	1554	1	US-08-045-269C-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-039-198A-1  
; Sequence 1, Application US/09039198A  
; Patent No. 6200951  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/039,198A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/34391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1636 base pairs  
; TYPE: nucleic acid

STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 2..1399  
 FEATURE:  
 NAME/KEY: mat peptide  
 LOCATION: 65..1399  
 US-09-039-198A-1

## Alignment Scores:

Pred. No.: 3,596-238 Length: 1636  
 Score: 2398.00 Matches: 445  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-039-198A-1 (1-1636)

QY 1 AlaIysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
 DB 65 GCAAACTGGTCTGCTACTTACCAACTGGGCCAGTACAGAGGGGGAGGCTCGCTTC 124  
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
 DB 125 CTGCCCNAGACTTGGACCCCGCTTGGACCCACTCATCTACGCCCTCGTGGCAATG 184  
 QY 41 ThrAsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
 DB 185 ACCAACCCAGCTGAGCACCCTGAGTGGAAATGACGAGACTCTTACACAGGAGTTCAAT 244  
 QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPhe 80  
 DB 245 GGCTTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTACCATCGAGGCTGGAAATTC 304  
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
 DB 305 GGCACTCAGAGTTTCAGATATGTGTAGCCAGCCCNACAAACCGTCAGACCTTTGTCAAC 364  
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
 DB 365 TCGGCCATCAGGTTCTGCGCAAAATACAGCTTTGACGGCTTACCTTGACTGGAGTAC 424  
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
 DB 425 CCAGGAAGCCAGGGAGCCCTGCGGTAGACAGGAGCGCTTCAACACCTGGTACAGGAC 484  
 QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160  
 DB 485 TTGGCCAAATGCTTCCAGCAGGAGGCCAGCCCTCAGGAGGAGACCGCTTCTTCTAGT 544  
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
 DB 545 GCAGCGGTTCCAGCTGGCAGAGCCTATGTGATGTGTGATACGAGTGGACAAATCGCC 604  
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
 DB 605 CAGAACTGGATTGTGCAACCTTATGGCTTACGACTTCATGGCTTTGGGAGAGGTC 664  
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaLeuSerLeu 220  
 DB 665 ACGGGACATACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGACGAGCCAGCCCTC 724  
 QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240  
 DB 725 AACGTGGATGCTCTGTGCAACAGTGGCTGCAAGAGGAGCCCTGCCAGAGCTGATC 784  
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
 DB 785 CTTGGCATGCTTACCTACGAGACGCTCTTACACTGGCTCTCATCATCAGACACGAGTG 844

QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
 DB 845 GGGGCCCCAGCCACAGGCTCTGCACCTCCAGGCCCTTCAACAGGAGGAGGATGCTG 904  
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
 DB 905 GCGTACTATGAAGTCTGCTCTCGAAGGGGGCCACCAACAGAGAAATCCAGGATCAGAAG 964  
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
 DB 965 GTGCCCTCATCTTCGGGACCAACAGTGGTGGGCTTTGATGATGGAGAGCTTCAAA 1024  
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyValMetValTrpAlaLeuAsp 340  
 DB 1025 ACCAAGTTCAGCTATCTGAAGCAGAGGAGCTGGGGGGGCCATGGTCTGGGCACCTGGAC 1084  
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
 DB 1085 TTAGATGACTTTGGCGGCTTCTCTCAACAGGCGCGATACCCCTCATCCAGACGCTA 1144  
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
 DB 1145 CGGCAGAACTCAGTCTTCCATCTTCCATCTTCCAGCACCCAGAGCTTGAAGTTCCAAA 1204  
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
 DB 1205 CCAGGTCAGCCCTCTGAACCTGAGCATGGCCCGCCCTGGACACAGACAGCTTCTGCCAG 1264  
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
 DB 1265 GSCAAAGCTGATGGGCTTATCCCAATCTCTCGGAACCGTCCAGCTTCTACAGCTGTGCA 1324  
 QY 421 AlaGlyArgLeuPheGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
 DB 1325 GCGGGGCGGCTGTTCAGCAAGAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA 1384  
 QY 441 CysCysThrTrpAsn 445  
 DB 1385 TGCTGCACCTGGAAT 1399

RESULT 2  
 US-08-877-599-1  
 ; Sequence 1, Application US/08877599  
 ; Patent No. 6372212  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gray, Patrick W.  
 ; TITLE OF INVENTION: Chitinase Materials and Methods  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/877,599  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/663,618  
 ; FILING DATE: 14-JUN-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Rin-Laures, Li-Hsien  
 ; REGISTRATION NUMBER: 33,547  
 ; REFERENCE/DOCKET NUMBER: 27866/33994  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1636 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..1399  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 65..1399  
US-08-877-599-1

## Alignment Scores:

Pred. No.: 3,596-238 Length: 1636  
Score: 2398.00 Matches: 445  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-08-877-599-1 (1-1636)

QY 1 AlaLysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
DB 65 GCAAAACTGGTCTGCTACTTCCAACTGGGCCCACTACAGACAGGGGAGGCTCGCTTC 124  
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
DB 125 CTGCCCAAGGACTTGGACCCCAAGCTTTGCACCCCACTCATCTACGCCCTTCCTGGCATG 184  
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
DB 185 ACCAACCCAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGGAGTTCAAT 244  
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPhe 80  
DB 245 GGCCTGAAGAAGTGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGCTGGAAATTC 304  
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
DB 305 GGCACCTCAGAGTTTCACAGATATGTAGCACCGCCACACCGTCAGACCTTTGTCAAC 364  
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
DB 365 TCGGCCCATCAGGTTTCTCGGCAAAATACAGCTTTGACGGCTTGACCTTGAGTGGAGTAC 424  
QY 121 ProGlySerGlnGlySerProAlaValAlaLysLysGluArgPheThrThrLeuValGlnAsp 140  
DB 425 CCAGGAAGCCAGGGAGCCCTCGCGTAGACAAGGAGCGCTTCACACCTCGTACAGGAC 484  
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160  
DB 485 TTGGCCATATGCTTCCAGCAGGAGCCCAAGACTCAGGAAGAGACGCTTCTCTGAGT 544  
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
DB 545 GCAGCGGTTCCAGCTGGGCAGACCTATGTGGATGTGGATACGAGGTGGACAAATCGCC 604  
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
DB 605 CAGAACTCGATTTTGTCAACCTTATGGCTACGACTTCATGGCTTTGGGAGAGGTC 664  
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluLysGlyValAlaAlaSerLeu 220  
DB 665 ACGGGACATAACAGCCCTCTTCAAGAGCGCAGAGAGAGTGTGCAGACCCGCTC 724  
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240

DB 725 AACGTGATGCTGCTGTGCACACAGTGGCTGCAGAAGGGGACCCCTGCAGCAAGCTGATC 784  
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
DB 785 CTGGCATGCTTACTCTACGAGCGCTCTTCACTTGGCTCTCTATCAGACACCAAGATG 844  
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
DB 845 GGGGCCCCAGCCACAGGCTTGGCACTCCAGGCCCTTCCACCAAGGAAGGAGGATGCTG 904  
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
DB 905 GCCTACTATGAAGTCTGCTCTCGAAGGGGGCCACCAACACAGAGAAATCCAGGATCAGAAG 964  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320  
DB 965 GTGCCCTACATCTTCCGGGACAAACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAAA 1024  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
DB 1025 ACCAAGGTCACTATCTGAAGCAGAAGGAGCTGGGGCGGGCCATGCTCTGGGCACTGGAG 1084  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
DB 1085 TTAGATGACTTTGGCGGCTTCTCTCTGCAACACAGGCGCATACCCCTCATCCAGACGCTA 1144  
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
DB 1145 CGGCAGGAACCTGAGTCTTCCATACACTTGGCTTTCAGGACCCCAAGAGCTTGAAGTTCCAAA 1204  
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
DB 1205 CCAGGTCAAGCTCTGAACCTGAGCATGGCCCAAGGCTGGGCAAGACACGCTTCTGCCAG 1264  
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
DB 1265 GGCAGAGCTGATGGGCTCTATCCCAATCTCTGGGAACGCTCCAGCTTCTACAGCTGTGCA 1324  
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
DB 1325 GCGGGCGGCTGTTCAGCAAGAGCTGCCGACAGGCTTGGTGTTCAGCAACTCTCTGCAAA 1384  
QY 441 CysCysThrTrpAsn 445  
DB 1385 TGCTGCACCTGGAAT 1399

## RESULT 3

US-09-267-574-1  
; Sequence 1, Application US/09267574  
; Patent No. 6399571  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE  
; FILE REFERENCE: 27866/35407  
; CURRENT APPLICATION NUMBER: US/09/267,574  
; CURRENT FILING DATE: 1999-03-12  
; EARLIER APPLICATION NUMBER: 09/039,198  
; EARLIER FILING DATE: 1998-03-12  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1636  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2) .. (1399)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (65) .. (1399)  
US-09-267-574-1

## Alignment Scores:

Pred. No.: 3 59e-238 Length: 1636  
Score: 2398.00 Matches: 445  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-267-574-1 (1-1636)

```
QY 1 AlaLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
DB 65 GCAAACTGGTCTGCTACTTCAACCACTGGGCCAGGTACAGACAGGGGAGGCTGCTTC 124
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 125 CTGCCAAGAGCTTGGACCCAGCCCTTTGACCACCTCATCTACGCCCTTCGCTGGCATG 184
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 185 ACCAACCCAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACCGAGGATTCAT 244
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 245 GGCCCTGAAGAAGATGAATCCCAAGCTGAAGACCCTGTAGCCATCGAGGCTGGAATTC 304
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 305 GGCACTCAGAAGTTTCAGATATGGTAGCCACCGGCCAACCAACCGTCAGACCTTTGTCAAC 364
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 365 TCGCCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCTTGACCTTGAGTGGAGTAC 424
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 425 CCAGGAAGCCAGGGAGCCCTGCGGTAGACAGAGGCGCTTCAACACCTCTGGTACAGGAC 484
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB 485 TTGGCCAAATGCCCTTCAGCAGGAAGCCAGACCTCAGGAAGGAGCCCTTCTCTGAGT 544
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 545 GCAGCGGTTCCAGCTGGCAGACCTATGTGATGCTGTGATACAGGTGACAAATCGCC 604
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 605 CAGAACCTGGATTTGTCACACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTC 664
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluLysSerGlyValAlaAlaSerLeu 220
DB 665 ACGGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGACGACGCCCTC 724
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
DB 725 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGACCCCTGCCAGCAAGCTGATC 784
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB 785 CTTGGCATGCTACCTACGAGCAGCTCCTTACACAGTGGCCCTCTCATCAGACACAGAGTG 844
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlnGlyMetLeu 280
DB 845 GGGGCCCCAGCCACAGGGCTGGCATCTCAGGCCCTTACCAGGAAGAGGGAGTGGCTG 904
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 905 GCCTACTATGAGTCTGCTCTCGAAGGGGGCCCAACCAAGAGATCCAGGATCAGAAG 964
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
DB 965 GTGCCCTATCTTCGGGACCAACCAAGTGGTGGGCTTTGATGATGTGAGAGCTTCAAA 1024
```

```
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyValMetValTrpAlaLeuAsp 340
DB 1025 ACCAAGGTGAGTACTTGAAGCAGAGGAGCTGGGGGGGCCCATGGTCTGGGCACTGGAC 1084
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCATACCCCTCATCCAGAGCGTA 1144
QY 361 ArgGlnGluLeuSerLeuProTyrIleuProSerGlyThrProGluLeuGluValProLys 380
DB 1145 CGCGCAGGAACCTGAGTCTTCCATACCTTGCAGCACCCAGAGCTTGAAGTTCACAAA 1204
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1205 CCAGGTGACCCCTCTGAACCTGAGCATGGCCCGCCCTGGACAGACAGCTTCTGCCAG 1264
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
DB 1265 GGCAAGCTGATGGGCTCTATCCCAATCCTCGGAACGGTCCAGCTTCTACAGCTGTGCA 1324
QY 441 CysCysThrTrpAsn 445
DB 1385 TGCTGCACCTGGAAT 1399
```

## RESULT 4

US-08-486-839-3  
; Sequence 3, Application US/08486839  
; Patent No. 5928928  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A human chitinase, its recombinant  
; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron  
; CITY: Jericho  
; STREET: 350 Jericho Turnpike  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11758  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486.839  
; FILING DATE: 07 - June - 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baron, Ronald J.  
; REGISTRATION NUMBER: 29,281  
; REFERENCE/DOCKET NUMBER: 294-26  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1643 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
US-08-486-839-3

Alignment Scores:

Pred. No.: 9,37e-238 Length: 1643  
Score: 2394.00 Matches: 444  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.83% Indels: 0  
DB: 2 Gaps: 0

US-10-004-219B-10 (1-445) x US-08-486-839-3 (1-1643)

```
QY 2 LysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AAACGTGGTCTGCTACTTCCCAACCTGGGCCAGTAGTACAGACAGGGGAGGCTCGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
DB 139 CCCAAGGACTTGGACCCAGCCCTTGGACCCACCTCATCTAGCCCTTCGCTGGCATGACC 198
QY 42 AsnHisGlnLeuSerThrTrpGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 199 AACCAACAGCTGAGCACCACCTGAGTGAATGACGAGACTCTCTACAGGAGTTCATATGCC 258
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGGAGGCTGGAAATTTCCGC 318
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 319 ACTCAGAAGTTTACAGATATGTTAGCCAGCCGACCAACACCTTGTCAACTCG 378
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTTCTCGCAAAATACAGCTTGTACGGCCTTGACCTTGAGTGGAGTACCCA 438
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 439 GGAAGCCAGGGGAGCCCTGCGCTAGCAAGGAGCGCTTCAACACCTTGTGACAGACTG 498
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 499 GCCATGCTTCCAGAGGAGCCAGACCTCAGGAGGAGACGCTTCTTCTGAGTGCA 558
QY 162 AlaValProAlaGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 559 GCGGTTCCAGCTGGGCAGACCTATGTGGATGCTGATACGAGTGGACAAATGCCAG 618
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 619 AACCTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGTCAAG 678
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyAlaAlaSerLeuAsn 221
DB 679 GGACATACAGCCCTTACAGAGGCGCAGAGAGAGTGGTGACAGCCAGCTCAAC 738
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATCCT 798
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
DB 799 GGCATGGCTTACCTACGAGCGCTTCTTCACTGGCTTCTCATCAGACACCAAGAGTGGGG 858
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281
DB 859 GCCCAGCCAGCAGGCTTGGCACTCAGGCCCTTCAACAGAGAGGAGGATGTGGCC 918
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
DB 919 TACTATGAATCTGCTCTCTGGAAGGGGGCCCAACAGAGAGATCCAGGATCAGAAGGTG 978
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321
DB 979 CCCTACATCTTCCGGACCAACAGTGGGTGGGCTTTGATGATGGAGAGCTTCAAAACC 1038
```

```
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341
DB 1039 AAGGTTCAGCTATCTGAGCAGAGGAGCTGGGGGGCATGGTCTGGGCACTGGACTTA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCGGCTTCTCTCTCAACACAGGGCGGATACCCCTCATCCAGACGCTACGG 1158
QY 362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381
DB 1159 CAGGACTGAGTCTTCCATACTTGGCTTCAGACACCCAGAGCTTGAAGTTCAAAACCA 1218
QY 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401
DB 1219 GGTCAAGCCCTCTGAACCTGAGCAGTGGCCCGCCAGCCCTGGACAAGACACGCTTCTGCCAGGGC 1278
QY 402 LysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAlaAla 421
DB 1279 AAAGCTGATGGGCTCTATCCCAATCTCTCGGAACGGTCCAGCTTCTACAGCTGTGCAGCG 1338
QY 422 GlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441
DB 1339 GGGCGGCTGTTCCAGCAAAAGCTGCCGACAGCCCTGGTGTTCAGCAACTCTCTGCAAAATGC 1398
QY 442 CysThrTrpAsn 445
DB 1399 TGCACCTGGAAT 1410

RESULT 5
US-09-151-011-3
; Sequence 3, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-09-151-011-3

Alignment Scores: 9,37e-238 Length: 1643
Pred. No.: 1643
```

Score: 2394.00 Matches: 444  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.83% Indels: 0  
DB: 3 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-151-011-3 (1-1643)

QY 2 LysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
DB 79 AAACCTGGTCTGCTACTTACCAACTGGGCCCGAGTACAGACAGGGGAGGCTGCTTCCTG 138  
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
DB 139 CCCAAGGACTGGACCCCGAGCCCTTGCACCCACCTCATCTACGCCCTTGCCTGGCATGCC 198  
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
DB 199 AACCCACGAGTACAGCACCTGAGTGGAAATGACGAGACTCTTACCCAGGAGTTCAATGGC 258  
QY 62 LeuLysLeuMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTTTAGCCATCGAGGCTGGAAATTTGGGC 318  
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
DB 319 ACTCAGAAGTTACACATATGTTAGTCCACCGGCCCAACACCGTCAGACCTTTGTCAACTCG 378  
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121  
DB 379 GCCATCAGGTTTCTGGCGAAATACACAGCTTTCAGCGCTTTCAGCTTGGGAGTACCCA 438  
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
DB 439 GGAAGCCAGGGAGCCCTCGCTAGACAGAGCGCTTTCACACCTGTTGACAGGACTTG 498  
QY 142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
DB 499 GCCAATGCCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAACGCTTCTTCTGAGTGCA 558  
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
DB 559 GCGGTTCCAGCTGGGAGAGCTATGTTGGATGCTGGATACAGGTGGACAAATCGCCAG 618  
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201  
DB 619 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTTCTGGAGAGGTCAGC 678  
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeuAsn 221  
DB 679 GGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGACAGCCGCTCAAC 738  
QY 222 ValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241  
DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGAGCCCTGCCAGCAAGCTGATCCTT 798  
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261  
DB 799 GGCATGCCCTTACCTACGAGCGCTCTTCACACTGGCTCTCATCAGACACAGAGTGGGG 858  
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281  
DB 859 GCCCCAGCCACAGGCTGGCACTCCAGGCCCTTCCACCAAGGAAGAGGAGTGGCTGGCC 918  
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301  
DB 919 TACTATGATGCTGCTGCTGGAAAGGGGGGCCACCAACAGAGATCCAGGATCAGAAAGGTG 978  
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321  
DB 979 CCTCATCTTCCGGACAAACAGTGGTGGCTTTGATGATGGAGAGCTTCAAAACC 1038  
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341

DB 1039 AAGGTGAGCTATCTCAAGCAGAAAGGAGTGGGGCGGGCCATGGTCTGGGCACTGGACTTA 1098  
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361  
DB 1099 GATGACTTTGGCGGCTTCTCTGCAACAGGCGCCATACCCCTCATCCAGACGCTACGG 1158  
QY 362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381  
DB 1159 CAGGAACCTGAGTCTTCCATACCTTTCAGGACCCCGAGAGCTTGAAGTTCCAAACCA 1218  
QY 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401  
DB 1219 GGTACAGCCCTCTGAACCTGAGCATGGCCCGCCAGCCCTGGACACACGCTTCTGCCAGGC 1278  
QY 402 LysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAlaAla 421  
DB 1279 AAAGCTGATGGGCTCTATCCCAATCTCGGACCGTTCAGCTTCTACAGCTGTGACGG 1338  
QY 422 GlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441  
DB 1339 GGGCGGCTGTTCCAGCAAGCTGCCCGACAGGCTGGTTCAGCAACTCTCTGCAATGC 1398  
QY 442 CysThrTrpAsn 445  
DB 1399 TGCACCTGGAAT 1410

## RESULT 6

US-09-343-623-3  
; Sequence 3, Application US/09343623  
; Patent No. 6303118

## GENERAL INFORMATION:

APPLICANT: A human chitinase, its recombinant  
TITLE OF INVENTION: production, its use for decomposing chitin, its use  
TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:

ADDRESSER: Hoffmann & Baron  
STREET: 350 Jericho Turnpike  
CITY: Jericho

STATE: New York

COUNTRY: United States of America

ZIP: 11758

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

SOFTWARE:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/343,623

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/486,839

FILING DATE: 07-June-1995

ATTORNEY/AGENT INFORMATION:

NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1643 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

US-09-343-623-3



## Alignment Scores:

Pred. No.: 9, 37e-238 Length: 1643  
Score: 2394.00 Matches: 444  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.83% Indels: 0  
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-343-623-3 (1-1643)

```
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AAACGGTCTGCTACTTACCAACATGGGCCCAAGTACAGACAGGGGGAGCTGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41
DB 139 CCCAAGGACTTGGACCCAGCCCTTGGACCCACCTCATCTACGCCCTTCTGCTGGCATGACC 198
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 199 AACCAACAGCTGAGCACCACTGAGTGGGAATGACGAGACTCTCTACCCAGAGTTCAATGGC 258
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyClyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGGAGGCTGGAAATTTCCGC 318
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 319 ACTCAGAAGTTCACAGATATGTTAGCCAGCGGCCCAACACCGTCAGACCTTTGTCAACTCG 378
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTCTCGGCAATACAGCTTTGACGGCCCTTGACCTTGAGTGGAGTACCCA 438
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 439 GGAAGCCAGGGAGGCCCTCGCCGTAGCAAGGAGCGCTTCACAAACCCTGGTACAGACTTG 498
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 499 GCCAATGCTTCCAGCAGGAGCCACGACCTCAGGAGAGGAAAGCGCTTCTCTGAGTGCA 558
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 559 GCGGTTCCAGCTGGGACAGACCTATGTGATGCTGATACGAGGTGCGACAAATGCCCCAG 618
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 619 AACCTGGATTTTGTCAACCTTATGGCCTACGACTTCCACTGGCTCTTGGGAGAAGTCCAG 678
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeuAsn 221
DB 679 GGACATACAGCCCTCTACAGAGGCGAAGAGAGAGTGGTGCAGCAGCCAGCTTCAAC 738
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241
DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAAGGGAGCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
DB 799 GGCATGGCTTACTACGAGCGCTCTCTTCACTGGCCTCTCATCAGACACCAAGAGTGGGG 858
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281
DB 859 GCCCCAGGCACAGGGTCTGGCACTCCAGGCCCTTCCAGAGAGGAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
DB 919 TACTATGAAGTCTGCTCTCTGGAAGGGGGCCCAACACAGAGATCCAGGATCAGAGGTG 978
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321
DB 979 CCCTACATCTTCCGGGACAACCAAGTGGGTGGGCTTTGATGATGGAGAGCTTCAAAACC 1038
```

```
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341
DB 1039 AAGGTACGTATCTGAAGCAGAGGGACTGGCGGGGCCATGGTCTGGGCACGTGACTTA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCGGCTTCTCTGCAACCCAGGGCCGATACCCCTCATCCAGACGCTACGG 1158
QY 362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381
DB 1159 CAGGAACCTGAGTCTTCCATCTTGCCTTTCAGGACCCCGAGAGCTTGAAGTTCCAAAACCA 1218
QY 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401
DB 1219 GGTCAAGCCCTCTGAACCTGAGCATGGCCCGACCCCTGGCAAGACACACCTTCTGCCAGGGC 1278
QY 402 LysAlaAspGlyLeuTyrProAsnProArgGluArgSerPheThrTyrSerCysAlaAla 421
DB 1279 AAGCTGATGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCAGCG 1338
QY 422 GlyArgLeuPheGlnGlnSerCysProThrThrGlyLeuValPheSerAsnSerCysLysCys 441
DB 1339 GGGCGGCTGTTCAGCAAAAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAAATGC 1398
QY 442 CysThrTrpAsn 445
DB 1399 TGCACCTGGAAAT 1410

RESULT 7
US-09-039-198A-3
; Sequence 3, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 27..1424
; FEATURE:
; NAME/KEY: mat_peptide
```

LOCATION: 90..1424

US-09-039-198A-3

## Alignment Scores:

Pred. No.:	1,53e-237	Length:	1656
Score:	2392.00	Matches:	444
Percent Similarity:	99.78%	Conservative:	0
Best Local Similarity:	99.78%	Mismatches:	1
Query Match:	99.75%	Indels:	0
DB:	4	Gaps:	0

US-10-004-219B-10 (1-445) x US-09-039-198A-3 (1-1656)

QY	1	AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe	20
DB	90	GCARAACTGGTCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTGCTTC	149
QY	21	LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet	40
DB	150	CTGCCCAAGGACTTGGACCCAGCCTTTGCACCCACTCATCTACGCCCTTCGCTGGCATG	209
QY	41	ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn	60
DB	210	ACCAACCCAGCTGAGCACCCTGAGTGGATGGAATGACGAGACTCTTACCCAGGAGTTCAAT	269
QY	61	GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe	80
DB	270	GGCTGGAAGAGATGATCCCAAGCTGAGACCCCTGTAGCCCATCGAGGCTGGAAATTC	329
QY	81	GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn	100
DB	330	AGCACTCAGAAGTTTACAGATATGTTAGCCAGCCGACCAACCAACCCCTGAGACCTTTGTCAAC	389
QY	101	SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr	120
DB	390	TCGGCCATCAGGTTCTTGGCAATACAGCTTTCAGCGGCTTGACCTTACCTGGGAGTAC	449
QY	121	ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp	140
DB	450	CCAGGAGCCAGGGGAGCCCTGCGGTAGACAGGAGCGCTTACACACCTTGGTACAGGAC	509
QY	141	LeuAlaAsnAlaPheGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSer	160
DB	510	TTGGCCATGCTTCCAGCAGGAGGCCAGACCTCAGGGAAGAAAGCGCTTCTTCTGAGT	569
QY	161	AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla	180
DB	570	GCAGCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACAGGTGGACAAAATCGCC	629
QY	181	GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal	200
DB	630	CAGAACTGGATTGTTCACCTTATGGCTAGACTTCCATGGCTCTTGGGAGAGGTC	689
QY	201	ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu	220
DB	690	ACGGGACATAACAGCCCTCTACAAGAGCAGAGAGAGAGTGGTGCAGCAGCCAGCTTC	749
QY	221	AsnValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProHisLysLeuIle	240
DB	750	AACGTGGATGCTGCTGCAACAGTGGCTGCAAGAGGGAGCCCTCCAGCAAGCTGATC	809
QY	241	LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal	260
DB	810	CTTGGCATGCTTACCTACGAGCGCTCTTCACTGAGCTTCCATCAGACACAGAGTG	869
QY	261	GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu	280
DB	870	GGGGCCCCAGCCAGGGCTTGGCACTCCAGGCCCTTCAACAAAGAGAGGAGGATGCTG	929
QY	281	AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys	300
DB	930	GGCTACTATGAAGTCTGCTCTTGGAGGGGGGCCACCAACAGAGAAATCCAGGATCAGAAG	989

QY	301	ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys	320
DB	990	GTGCCCTACATCTTCGGGCAACACCACTGGGTGGCTTTGATGATGTGGAGAGCTTCAAA	1049
QY	321	ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp	340
DB	1050	ACCAAGGTCAGCTATCTGAAGCAGAAGGAGTGGGGCGGGGCCCATGGTCTGGGCACTGGAC	1109
QY	341	LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu	360
DB	1110	TTAGATGACTTTCGGGCTTCTCCGCAACAGGCGCCATACCCCTCATCCAGAGGCTA	1169
QY	361	ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys	380
DB	1170	CGGCAGGAACCTGAGTCTTCATCTTCCAGGACCCAGAGCTTCAAGTTCACAAA	1229
QY	381	ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln	400
DB	1230	CCAGGTACGCCCTCTGAACCTGAGCATGGCCCGCCAGGCGGCAAGACACGTTCTGCCAG	1289
QY	401	GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla	420
DB	1290	GGCAAGCTGATGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGA	1349
QY	421	AlaGlyArgLeuPheGlnLysSerCysProThrGlyLeuValPheSerAsnSerCysLys	440
DB	1350	GGGGGGCGCTTCTCCAGCAAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA	1409
QY	441	CysCysThrTrpAsn	445
DB	1410	TGCTGCACCTGGAAT	1424

RESULT 8  
US-08-877-599-3  
; Sequence 3, Application US/08877599  
; Patent No. 6372212  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; TITLE OF INVENTION: Chitinase Materials and Methods  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/877,599  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/663,618  
; FILING DATE: 14-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/33994  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1656 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 27..1424  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 90..1424  
US-08-877-599-3

## Alignment Scores:

Pred. No.: 1,536-237 Length: 1656  
Score: 2392.00 Matches: 444  
Percent Similarity: 99.78% Conservative: 0  
Best Local Similarity: 99.78% Mismatches: 1  
Query Match: 99.75% Indels: 0  
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-08-877-599-3 (1-1656)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyAlaAlaArgPhe 20  
DB 90 GCAAAACTGGTCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTGGCTTC 149  
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
DB 150 CTGCCCAAGGACTTGGACCCAGCCTTTGCACCACCTCATCTACGCTTTCGCTGGCATG 209  
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
DB 210 ACCAACCCAGCTGAGCACCCTGAGTGGGATGACGAGACTCTTACCAGGAGTTCAAT 269  
QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80  
DB 270 GGCCTGAAGAGATGAATCCCAAGCTGAAGACCTCTGTAGCCATCGAGGCTGGAATTC 329  
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
DB 330 AGCACTCAGAAGTTACAGATATGTTAGCCAGCCGACCAACCGTCAGACCTTTGTCAAC 389  
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
DB 390 TGGCCATCAGGTTCTGCGCAATACAGCTTTCAGGCTTTCAGCTTTCAGCTGGGAGTAC 449  
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
DB 450 CCAGGAAGCCAGGGAGCCCTGCGTAGACAGGAGCGCTTCAACCTCTGGTACAGGAC 509  
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
DB 510 TTGGCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGGAGACGCTTCTTCTGAGT 569  
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
DB 570 GCAGCGGTTCAGCTGGGAGAGCTATGGATGCTGATACAGGTGACAAATCGCC 629  
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
DB 630 CAGAACTGGATTGTTCAACCTTATGGCTACAGCTTCCATGGCTCTTGGGAGAGGTC 689  
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu 220  
DB 690 ACGGACATAACAGCCCTCTACAAGAGGCAAGAGAGTGGTCAGCAGCAGCCCTC 749  
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240  
DB 750 ACGTGGATGCTGCTGTGCAACAGTGGCTGCAAGGGGACCCCTGCCAGAGCTGATC 809  
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
DB 810 CTTGGCATGCTTACTACGGAGCGCTCTTCCACACTGGCTCTCTCATCAGACACAGATG 869  
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280

DB 870 GGGGCCCCAGCCACAGGCTCTGGCACTCCAGGCCCTTCCAAAGAAAGAGGGATGCTG 929  
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
DB 930 GCCTACTATGAAGTCTGCTCTGGNAGGGGGCCCAACAGAGAGATCCAGGATCAGAAG 989  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
DB 990 GTGCCCTACATCTTCCGGGCAACCCAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1049  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
DB 1050 ACCAAGGTGAGCTATCTGAAGCAGAGGAGCTGGGGCGGGCCCATGGTCTGGGCACCTGGAC 1109  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
DB 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACCAAGGCCCATACCCCTCATCCAGAGCTA 1169  
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
DB 1170 CGGCAGGAAGTGAAGTCTTCCATCTTGGCTTTCAGGACCCAGAGCTTGAAGTTCCAAA 1229  
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
DB 1230 CCAGTCAAGCCCTTGAACCTGAGCATGGCCCGCCCTGGACAAGACACGTTCTGCCAG 1289  
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
DB 1290 GGCAGAGCTGATGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCA 1349  
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
DB 1350 GGGGGCGGCTGTTCAGCAAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA 1409  
QY 441 CysCysThrTrpAsn 445  
DB 1410 TGCTGCACCTGGAAT 1424

## RESULT 9

US-09-267-574-3  
Sequence 3, Application US/09267574  
Patent No. 6395571  
GENERAL INFORMATION:  
APPLICANT: Gray, Patrick W.  
APPLICANT: Tjoelker, Larry W.  
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
FILE REFERENCE: 27866/35407  
CURRENT APPLICATION NUMBER: US/09/267,574  
CURRENT FILING DATE: 1999-03-12  
EARLIER APPLICATION NUMBER: 09/039,198  
EARLIER FILING DATE: 1998-03-12  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 1656  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (27)..(1424)  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: (90)..(1424)  
US-09-267-574-3

Alignment Scores:  
Pred. No.: 1,536-237 Length: 1656  
Score: 2392.00 Matches: 444  
Percent Similarity: 99.78% Conservative: 0  
Best Local Similarity: 99.78% Mismatches: 1  
Query Match: 99.75% Indels: 0  
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-267-574-3 (1-1656)

QY 1 AlAlsLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20  
Db 90 GCAAACTGGTGTCTACTTCAACCACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTC 149  
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40  
Db 150 CTGCCCAAGGACTTGGACCCAGCCCTTGGACCCACCTCATCTACGGCTTCGCTGGCATG 209  
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
Db 210 ACCAACCCAGCTGAGCAGCAGCTAGTGGCAATACAGAGACTCTTACACGAGGATCAAT 269  
QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80  
Db 270 GGCTTGAAGAAGATGAATCCAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTC 329  
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
Db 330 AGCACTCAGAGTTTCCAGATATGGTAGCCACGGCCCAACACCGTCAGACTTTGTCTAAC 389  
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 390 TCGGCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGACTGGAGTAC 449  
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
Db 450 CCAGGAAGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCAACACCCCTGGTACAGGAC 509  
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160  
Db 510 TTGGCCATCCCTTCCAGCAGGAGGCCACACCTCAGGGAGGAAACGCTTCTTCTGAGT 569  
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
Db 570 GCAGCGGTTCAGCTGGCCAGACCTATGTGGATGTGGATACGAGGTGGACAAATCGCC 629  
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
Db 630 CAGAACCTGGATTTGTCAACCTTATGGCTAGACTTCCATGGCTCTTGGGAGAGGTC 689  
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaLaserLeu 220  
Db 690 ACGGACATACAGCCCTCTACAGAGGCAAGAGAGTGTGCAGCAGCCAGCCCTC 749  
QY 221 AsnValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProLaserLysLeuIle 240  
Db 750 AACGTGGATGCTGTGTCAACAGTGGCTGCAGAAAGGGACCCCTGCGCAGAGCTGATC 809  
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
Db 810 CTGGGATGCTTACCTACGAGCGCTCTTCACTAGCTGGCTTCCATCAGACACAGAGTG 869  
QY 261 GlyValaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280  
Db 870 GGCGCCACCCAGCGTCTGGCACTCCAGGCCCTTCCAGAGGAGGAGGAGGATGCTG 929  
QY 281 AlaTyrTrpGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 930 GCCTACTATGAGTCTGCTCTGGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 989  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320  
Db 990 GTGCCCTTACATCTTCCGGGCAACACCTAGTGGGTGGCTTTGATGTGGAGAGCTCAAA 1049  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340  
Db 1050 ACCAAGTCACTATCTGAAGCAGAGGAGTGGCGGGCCATGCTTGGGCACTGGAC 1109  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360  
Db 1110 TTAGATGACTTTTCCGGCTTCTCTCTGCAACAGGGGGCGATACCCCTCTATCCAGAGCTA 1169

QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380  
Db 1170 CGCGAGGAAGTCTGAGTCTTCCATCTTGCCTTCCAGCACCCAGAGCTTGAAGTTCCAAA 1229  
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
Db 1230 CCAGGTCAGCCCTCTGAACCTGAGCATGSCCCAGCCCTGGACAAGACACGCTTCTGCCAG 1289  
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
Db 1290 GGCAAAAGCTGATGGGCTCTATCCCAATCCTCGGGAACGCTCCAGCTTCTACAGCTGTGA 1349  
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440  
Db 1350 GCGGGGGCGCTGTTCCAGCAAGCTGCCGACAGCGCTGGTGTTCAGCAACTCTCTGCAAA 1409  
QY 441 CysCysThrTrpAsn 445  
Db 1410 TGCTGCACCTGGAAT 1424

RESULT 10  
US-08-486-839-5  
; Sequence 5, Application US/08486839  
; Patent No. 5928928  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A human chitinase, its recombinant  
; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11758  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,839  
; FILING DATE: 07 - June - 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baron, Ronald J.  
; REGISTRATION NUMBER: 29,281  
; REFERENCE/DOCKET NUMBER: 294-26  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1713 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; US-08-486-839-5

Alignment Scores:  
Pred. No.: 5,048-233 Length: 1713  
Score: 2348.50 Matches: 443  
Percent Similarity: 94.66% Conservative: 0  
Best Local Similarity: 94.66% Mismatches: 1  
Query Match: 97.94% Indels: 25  
DB: 2 Gaps: 1

US-10-004-219B-10 (1-445) x US-08-486-839-5 (1-1713)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
DB 79 AAACGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTCTCTG 138  
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
DB 139 CCCAAGGACTTGGACCCCAAGCCCTTTGCACCCACCTCATCTAGCCCTTGGCTGGCATGCC 198  
QY 42 AsnHisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
DB 199 AACCAACAGCTGAGCACCACCTGAGTGGATGAGAGACTCTCTACAGAGTTCAATGGC 258  
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPheGly 81  
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAAATTCGGC 318  
QY 82 ThrGlnLysPheThrAspMetValIleAlaAsnAsnArgGlnThrPheValAsnSer 101  
DB 319 ACTCAGAAGTTTACAGATATGTTAGCCACCGGCCAACCAACCGTCAGACCTTTGTCAACTCG 378  
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGlnTyrPro 121  
DB 379 GCCATCAGTTTCTGGCAATATACAGCTTTGACGGCTTGACCTTGACCTGGGAGTACCCA 438  
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
DB 439 GGAAGCCAGGGAGGAGCCCTGCGCTAGCAAGAGAGCGCTTACACACCTGGTACAGACTTG 498  
QY 142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
DB 499 GCCAATGCTTTCAGCAGGAAGCCAGACCTCAGGGAAGAAAGCGCTTCTTCTGAGTGCA 558  
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
DB 559 GGGGTTTCAGCTGGGAGACCTATGTGGATGTGTGATAGAGTGGACAAAATCGCCAG 618  
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201  
DB 619 AACCTGGATTTTGTCAACCTTATGGCTACGACTCCATGGCTCTTGGGAGAGGTCAAG 678  
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaSerLeuAsn 221  
DB 679 GGACATAACAGCCCTCTTACAAGAGGCAAGAGAGAGTGTGTGACAGCCAGCCAGCTCAAC 738  
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241  
DB 739 GTGGATGCTGTGTGCAACAGTGGCTGCAAGAGGGAGCCCTGCGCAGCAAGCTGATCCTT 798  
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGly 261  
DB 799 GGCATGCTTACCTACGGAGCTCTCTTACACTGGCTCTCTCATCAGACACAGAGTGGGG 858  
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlnGlyGlyMetLeuAla 281  
DB 859 GCGCCAGCCACAGGGTCTGGCACTCAGGCCCTTTCACCAAGAGAGGGAGTGTGGCC 918  
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301  
DB 919 TACTATGAAGTCTGCTCTGGGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAGGTG 978  
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321  
DB 979 CCTCATACCTTCCGGGACCAACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAAAACC 1038  
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341  
DB 1039 AAGGTCAAGTATCTGAAGCAGAAGGAGACTGGGGGGGGCCATGGCTGTGGGCACTGGACTTA 1098  
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361  
DB 1099 GATGACTTTGGCCGGCTTCTCTTGCACCAAGGGCCGATACCCCTCATCCAGACGCTACGG 1158

QY 362 GlnGluLeuSer----- 365  
DB 1159 CAGGAACCTGAA-TGGGTAAAGCCTTAACCTGCTGTACATGTGAGGCCAGGTGTTGCTGTG 1217  
QY 366 -----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377  
DB 1218 TGGCACTGTGCTTACAGCTGTAGGTCTTCCATACCTTTCAGGCACCCAGAGCTTGAA 1277  
QY 378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397  
DB 1278 GTTCCAAACACAGGTACGCTCTGAACTGAGCATGGCCCAAGCCCTGGACACAGACAG 1337  
QY 398 PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyr 417  
DB 1338 TTCTGCCAGGGCAAGCTGTAGTGGCTCTATCCCAATCTCGGGAACGCTCCAGCTTCTAC 1397  
QY 418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437  
DB 1398 AGCTGTGAGCGGGGGCGCTGTTCAGCAAAAGCTGCCCGCAGAGGCTGTGTTCAGCAAC 1457  
QY 438 SerCysLysCysCysThrTrpAsn 445  
DB 1458 TCCTGCAAAATGCTGCACCTGGAAT 1481  
RESULT 11  
US-09-151-011-5  
; Sequence 5, Application US/09151011  
; Patent No. 6057142  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant  
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in  
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Hoffmann & Baron, LLP  
; STREET: 6900 Jericho Turnpike  
; CITY: Syosset  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11791  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/151.011  
; FILING DATE: 10 - September - 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Morris, Robert C.  
; REGISTRATION NUMBER: 42,910  
; REFERENCE/DOCKET NUMBER: 294-32 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1713 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
US-09-151-011-5  
Alignment Scores:  
Pred. No.: 5, 04e-233 Length: 1713  
Score: 2348.50 Matches: 443  
Percent Similarity: 94.66% Conservative: 0  
Best Local Similarity: 94.66% Mismatches: 1  
Query Match: 97.94% Indels: 25

DB: 3 Gaps: 1

US-10-004-219B-10 (1-445) x US-09-151-011-5 (1-1713)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
DB 79 AAACCTGCTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCTCTG 138

QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
DB 139 CCCAAGGACTTGGACCCGACCTTTTGACCCACCTCATCTACGCCCTTGGCTGGCATGACC 198

QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
DB 199 AACCCACGCTGAGCACCCTGAGTGGATGACGAGACTCTCTACAGGAGTTCANTGGC 258

QY 62 LeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGAGGCTGGAAATTCGGC 318

QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
DB 319 ACTCAAGAGTTTCAGATATGGTAGCCACGCGCCCAACCGTCAGACCTTTGTCAACTCG 378

QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121  
DB 379 GCCATCAGGTTTCTGGCAAAATACAGCTTTGACGGCCTTGACCTTGAGTGGAGTACCCA 438

QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
DB 439 GGAAGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCAACACCTGGTGTACAGGACTTG 498

QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
DB 499 GCCAATGCCCTTCAGCAGGAGCCAGACCTCAGGGAAGAGCGCTTCTTCTGAGTGCA 558

QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
DB 559 GCGGTTCCAGCTGGCAGACCTATGTGGATGCTGGATACGAGGTGGACAAATCGCCAG 618

QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201  
DB 619 AACCTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTTGGGAGAGGTCACG 678

QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluLysSerGlyAlaAlaLaserLeuAsn 221  
DB 679 GGAATACAGCCCTCTTCAAGAGGCAAGAGAGTGGTGCAGCGCCAGCCCTCAAC 738

QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241  
DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798

QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261  
DB 799 GGCATGCTCTACCTACGAGCGCTCTTACACTGGCCCTCTCATCAGACACAGAGTGGG 858

QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281  
DB 859 GCCCCAGCCAGGGCTGGCATCTCAGGCCCTTCCACCAAGGAAGAGGAGTGGCTGGCC 918

QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301  
DB 919 TACTATGAGTCTGCTCTGGAAGGGGGCCCAACCAACAGAGAATCCAGGATCAGAAGGTG 978

QY 302 ProTyrIlePheArgAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321  
DB 979 CCTATCATCTTCCGGGCAACACCAAGTGGTGGCTTGTGATGATGGAGAGCTTCAAAACC 1038

QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341  
DB 1039 AAGGTGAGTATCTGAAGCAGAGGAGGACTGGGGGGGCCCATGCTTGGGCACTGGACTTA 1098

QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361

DB 1099 GATGACTTTGGCGGCTTCTCTCTGCAACAGGGCCGATATACCCCTCATCCAGACGCTACGG 1158

QY 362 GlnGluLeuSer-----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377

DB 1159 CAGGAACGTAA--TGGGTAAAGCCTTAACTGCTGTCACTTTCCTTTCAGGACCCAGAGCTTGA 1277

QY 366 -----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377

DB 1218 TGGCACTGTGCTTACGCTGTAGTCTTTCATACTTTCCTTTCAGGACCCAGAGCTTGA 1277

QY 378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397

DB 1278 GTTCCAAAACAGGTGAGCCCTCTCAACTGAGCATGGCCCAAGCCCTGGACAGACACG 1337

QY 398 PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerPheTyr 417

DB 1338 TTCTGCCAGGGGCAAGCTGATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTAC 1397

QY 418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437

DB 1398 AGCTGTGACGGGGGGGCTGTTCAGCAAGCTGCCCGACAGGCTGGTGTTCAGCAAC 1457

QY 438 SerCysLysCysCysThrTrpAsn 445

DB 1458 TCCTGCAAAATGCTGCACCTGGAAT 1481

RESULT 12

US-09-343-623-5  
; Sequence 5, Application US/09343623  
; Patent No. 6303118  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A human chitinase, its recombinant  
; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
; NUMBER OF SEQUENCES: 16  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Hoffmann & Baron  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11758  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; SOFTWARE:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/343,623  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,839  
; FILING DATE: 07-June-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baron, Ronald J.  
; REGISTRATION NUMBER: 29,281  
; REFERENCE/DOCKET NUMBER: 294-26  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1713 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; US-09-343-623-5

## Alignment Scores:

Pred. No.: 5,04e-233 Length: 1713  
Score: 2348.50 Matches: 443  
Percent Similarity: 94.66% Conservative: 0  
Best Local Similarity: 94.66% Mismatches: 1  
Query Match: 97.94% Indels: 25  
DB: 4 Gaps: 1

US-10-004-219B-10 (1-445) x US-09-343-623-5 (1-1713)

```
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AAACCTGGTCTGCTACTTCAACCACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41
DB 139 CCCAAGGACTTGACCCCGAGCCCTTTGCACCCACCTCATCTACGCCCTTCCTGGCATGACC 198
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 199 AACCAACCACTGAGCACCACCTGAGTGGATGACGAGACTCTCTACCAAGGAGTTCAATGGC 258
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTTAGCCATCGAGGCTGGAAATTCGGC 318
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 319 ACTAGAAGTTTACAGATATGTTAGCCAGCGGCCAACACCGTCAGACCTTTTGTCAACTCG 378
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 379 GCATCAGGTTTCTGCGCAATATACAGCTTTGACGGCTTGACCTTGAGTGGAGTACCCA 438
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 439 GGAAGCCAGGGAGCCCTGCGCTAGACAAAGAGCGCTTCACACCCCTGTTACAGACTTG 498
QY 142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 499 GCAATGCTTTCAGCAGGAAGCCAGACTCAGGGAAGAACGCTTCTTCTGTAGTGA 558
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 559 GCGGTTCCAGCTGGGCAGACCTATGTGGATGCTGGATACAGAGTGACAAATCGCCAG 618
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 619 AACCTGGATTTGTCAACCTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTACAG 678
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeuAsn 221
DB 679 GGACATAACAGCCCTCTACAGAGCGCAAGAGAGTGGTGACAGCCAGCCTCAAC 738
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
DB 739 GTGGATGCTGCTGCAACAGTGGCTGCAAGGGGACCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
DB 799 GGCATGCCCTACCTACGAGCGCTCTTTCACATGGCCCTCTCATCAGACCCAGAGTGGG 858
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281
DB 859 GCCCCAGCCACAGGGTCTGGCATCTCAGGCCCTTCCACCAAGGAAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
DB 919 TACTATGAAGTCTGCTCTCTGGGAAGGGGGCCACCAACAGAGATCAGAGTCAAGGTTG 978
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321
DB 979 CCTACATCTTCCGGGACCAACCACTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAAACC 1038
```

```
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341
DB 1039 AAGGTACGTATCTGNAAGCAGAGGACTGGCGGGCCATGGTCTGGGCACCTGACITTA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCGGCTTCTCCTGCAACCCAGGGCGGATACCCCTCATCCAGACGCTACGG 1158
QY 362 GlnGluLeuSer----- 365
DB 1159 CAGGAACCTGAA-TGGGTAAAGCCTTAACCTGCTGTCAATGTAGGCCAGGCTGTTCCTG 1217
QY 366 -----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377
DB 1218 TGGCAGCTGTGCTTCAGCTGTAGTCTTCCATACCTTCAGGCACCCAGAGCTTGAA 1277
QY 378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397
DB 1278 GTTCCAAACCCAGGTCAGCCCTCTGAACCTGAGCATGGCCCAAGCCCTGGACACACAGC 1337
QY 398 PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyr 417
DB 1338 TTCGCCAGGGCAAGCTGTATGGGCTCTATCCCAATCTCGGAAACGGTCCAGCTTCTAC 1397
QY 418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437
DB 1398 AGCTGTGAGCGGGGCGGCTGTTCAGCAAGAGCTGCCCGCAGAGGCTGGTGTTCAGCAAC 1457
QY 438 SerCysLysCysCysThrTrpAsn 445
DB 1458 TCCTGCAAAATGTCACCTGGAAT 1481
RESULT 13
US-08-694-915-1
; Sequence 1, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert.
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1433 base pairs
```

; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
US-08-694-915-1

## Alignment Scores:

Pred. No.: 1,33e-98 Length: 1433  
Score: 1046.50 Matches: 210  
Percent Similarity: 65.70% Conservatives: 85  
Best Local Similarity: 46.77% Mismatches: 117  
Query Match: 43.64% Indels: 38  
DB: 1 Gaps: 10

US-10-004-219B-10 (1-445) x US-08-694-915-1 (1-1433)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaAspPheLeu 21  
DB 118 AAAGTGGTTTCTACTTTACAACTGTCTCCAGACCGGAGAACACGAGAAATTCACC 177  
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
DB 178 CCTGAGAATATGACCCCTTCTCATGCTCTCATCTCATCTATTCAITTCGCGCAGCATCGAA 237  
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
DB 238 AACAAACAGGTTATCATCAAGGACAGAGTGAAGTGATGCTCTACCAAGACCATCAACAGT 297  
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81  
DB 298 CTCAAAACCAAGATCCCAACTGAAATTTCTCTTGCCATTGGAGGTACCTGTTTGGT 357  
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
DB 358 TCCAAAGGGTTCCACCTATGTTGGTATTTCTTACATCACGCTTGGAAATTCATTAACCTC 417  
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121  
DB 418 ATAATCTGTTCTGAGGAACCACTATTTGATGGACTGGATGTAAGCTGGATCAACCCA 477  
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
DB 478 GATCAGAAA-----GAAACACTCATTTCACTGTGCTGATCATGAGTTA 522  
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
DB 523 GCAGAGCCCTTCAGAGGACTTCACAAATTCACCAAGGAAAGGCTTCTCTTGAAGTGG 582  
QY 162 AlavalProAlaGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
DB 583 GGCATATCTCAGGGAGGCAAAATGATGTAACAGCTATCAAGTTGAGAAACTGGCAAAA 642  
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLys----- 199  
DB 643 GATCTGATTTTCAACACTCTCTGCTTTCATGCTTTCATGGGTCTTGGGAAAGCCCTT 702  
QY 200 ValThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAsp 219  
DB 703 ATCACTGGCCACAAACGCCCTCTGAGCAGGGGTGGCAGCAGAGGCCCAAGCTCCTAC 762  
QY 220 LeuAsnValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeu 239  
DB 763 TACAATGTGGAATATGCTGGGGGTACTGGATACATAAGGGAATGCCATCAGAGAAGGTG 822  
QY 240 IleLeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArg 259  
DB 823 GTCATGGGCATCCCATATGGGACTCTCTTACACTGGCC---TCTGCAGAAACACC 879  
QY 260 ValGlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMet 279

DB 880 GTGGGGGCCCTGCTCTGCGCTGGAGCTGTGGACCCATCACAGAGTCTTCAGGCTTC 939  
QY 280 LeuAlaTyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAsp 298  
DB 940 CTGGCCTATTATGAGATCGCAGTCTCTTAAAGATTTAAACCTGGGAGGAGCCATGATCTGGTCT 999  
QY 299 GlnLysValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSer 318  
DB 1000 CAGCAGTTCCCTACCGAGTCAAGGGGAACAGTGGTGGCTATCATCATGTGAAGACT 1059  
QY 319 PheLysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAla 338  
DB 1060 ATGGAGACCAAGTTCAGTCTCTTAAAGATTTAAACCTGGGAGGAGCCATGATCTGGTCT 1119  
QY 339 LeuAspLeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGln 358  
DB 1120 ATTGACATGGATGACTTCACTGGCAATCTGCAACACGAGGCCCTTACCTCTTGTCCA 1179  
QY 359 ThrLeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluVal 378  
DB 1180 GCAGTCAAGAGAGCCTT-----GGCTCCCTGTGAAGGATTAACCTTA 1221  
QY 379 ProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPhe 398  
DB 1222 CAGAGAAGCAGGCAAGATGACCTGTGCTGCTG-GGCCTGCTCTCTCCAGGAATTCATCA 1280  
QY 399 CysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSer 418  
DB 1281 TGT-----GGGATT-----CCCTTGCAGGATGGCCTTT----- 1310  
QY 419 CysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeu----- 433  
DB 1311 -----GGATCTCTCTC---CAAGCTTCTCTGACTTCTCTTAGATCATAGATTG 1358  
QY 434 -----ValPheSerAsnSerCys 439  
DB 1359 GACCTGGTTTCTGTTTCTCTGCGAGCTGT 1385  
RESULT 14  
US-08-694-915-3  
; Sequence 3, Application US/08694915  
; Patent No. 5811535  
; GENERAL INFORMATION:  
; APPLICANT: Adamou, Julie  
; APPLICANT: Kirkpatrick, Robert  
; APPLICANT: Rosenberg, Martin  
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-2799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/694,915  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Han, William T  
; REGISTRATION NUMBER: 34,344  
; REFERENCE/DOCKET NUMBER: ATG50017  
; TELECOMMUNICATION INFORMATION:



TELEPHONE: 610-270-5219  
TELEFAX: 610-270-5090  
TELEX.

INFORMATION FOR SEQ ID NO: 3:

### SEQUENCE CHARACTERISTICS:

LENGTH: 1526 base pairs

**TYPE:** nucleic 'acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO  
DOCUMENT BY DE

FRAGMENT TYPE:	ORIGINAL SOURCE:
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

ORIGINAL SOURCE:  
C94-91E-3

094-913-3

ent. Scores:

No. :

•

**Sir**

ocal

## Match

Length: 1526

Matches: 210

Conservative: 85

Mismatches: 117

Indels: 38

**Gaps: 10**

004-219B-10 (1-445) x US-08-694-915-3 (1-1526)

2 LysLeuValCysTyrPheThrAsnTrrAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
|||||  
211 AAACCTGGTTTGCCTACTTTACCAACTGGTCCAGAGCCGGCAGGAACCCAGGAATAATTCACC 270  
|||||  
22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
|||||  
271 CCTGAGAAATATTGACCCCTTCTATGCTCTCATCTCATCTATTTCATTCGCCAGCATCGAA 330  
|||||  
42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
|||||  
331 AACAAACAGGTTTATCATCAAGGACAAGAGTGAAGTATGCTCTTACCAGACCATCAACAGT 390  
|||||  
62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
|||||  
391 CTCAAAACCAAGAAATCCAAACTGAAATAATCTCTGTGCCAATTGGAGGGTACCTGTTTGT 450  
|||||  
82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAspArgGlnThrPheValAsnSer 101  
|||||  
451 TCCAAAGGGTTCCACCCTATGGTGGATCTCTTACATCAGCTTGGAAATTCATTAATCTCC 510  
|||||  
102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121  
|||||  
511 ATAATCTGTCTTCTGAGGAACCACTAACTTTGATGGACTGGATGAAGCTGGATCTACCCA 570  
|||||  
122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
|||||  
571 GATCAGAAA-----GAAACCACTATTTCATCTGTGCTGATTTCATGAGTTA 615  
|||||  
142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAla 161  
|||||  
616 GCAGAAACCTTTCAGAGGACTTCACAAATTCACCAGGAAGGCTTCTCTGACTGGCG 675  
|||||  
162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
|||||  
676 GGCGTATCTCGAGGGAGGCAAAATGATTGATAACAGCTATCAAGTTCAGAAATCGGCAAAA 735  
|||||  
182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLys----- 199  
|||||  
736 GATCTGATTTCATCAACCTCCCTGTCTCTTGGATTCATGGGTCTTGGGAAAGACCCCTT 795  
|||||  
200 ValThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSer 219  
|||||  
796 ATCACTGCCACACACCCCTCTGAGCAAGGGGTGGCAGACAGAGGGCCAGCTCTCTAC 855  
|||||  
220 LeuAsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 239  
|||||  
856 TACAATCTGGAATATGCTGTGGGTACTGATCATATAAGGGAATGCATCAGAGAAGGTG 915

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/455-5100  
TELEFAX: 619/455-5110  
INFORMATION FOR SEO ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1681 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: YKL-40  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 135..1681  
PCT-US94-07754-4

## Alignment Scores:

Pred. No.: 8.69e-97 Length: 1681  
Score: 1030.00 Matches: 202  
Percent Similarity: 67.33% Conservative: 68  
Best Local Similarity: 50.37% Mismatches: 115  
Query Match: 42.95% Indels: 16  
DB: 5 Gaps: 8

US-10-004-219B-10 (1-445) x PCT-US94-07754-4 (1-1681)

QY 2 LysLeuValCysThrPheThrAsnTTPAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
DB 138 AAATGCTGTGCTACTACACCATGCTGCTCCAGTACCGGAGGCGATGGAGCTGCTC 197  
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
DB 198 CCAGATGCCCTTCACCGCTTCCTGTGTATACCATCATCATCTACAGCTTGCCAAATAGC 257  
QY 42 AsnHisGlnLeuSerThrThrGluTTPAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
DB 258 AACGATCATATCGACACCTCGGAGTGGATGATGTACGCTCTACGCGATGCTCAACACA 317  
QY 62 LeuLysIleMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTTPAsnPheGly 81  
DB 318 CTCAACACACGAAACCCCAACCTGAGACTCTCTGTGTCGGAGATGGAATTTGGG 377  
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
DB 378 TCTCAAGATTTTCCAAGATAGCTTCCACACCCAGAGTCGCGGACTTTCATCAAGTCA 437  
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspThrGluTyrPro 121  
DB 438 GTACCGCCATTTCTGGCGACCCCATGGCTTTGATGGCGGTGACCTTGCTGCTTACCT 497  
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
DB 498 GGACGGAGA-----GACAAACACCATTTTACCACTTTCACCGCTTAATCAAGGAATG 542  
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
DB 543 AAGCGCGAATTTATAAGGAAGCCAG---CCAGGGAAGAAAGCAGCTCCTGCTCAGCGCA 599  
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
DB 600 GCACCTGTCTGGGGGAGGTGTCACCATGACAGAGCTATGACATTTGCCAAGATATCCCAA 659  
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerThrGluLysValThr 201  
DB 660 CACCTGGATTTTATAGCATCATGACCTACGATTTTCATGGCGCTTGGCGTGGACCA 719  
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluLysSerGlyAlaAlaSerLeuAsn 221  
DB 720 GGCCATCACAGTCCCTGTTCCGAGGTGAGGAGATGCAAGTCTCACAGATTCAGCAAC 779  
QY 222 ValAspAlaValGlnGlnThrProLysGlyThrProAlaSerLysLeuIleLeu 241

DB 780 ACTGACTATGCTGTGGGTGATCATGTTGAGGCTGGGGGCTCTCCAGTAAGCTGTGATG 839  
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261  
DB 840 GGCATCCCACTTCGGAGAGCTTCACTCTGGGT---TCTTCTGAGACTGTTGTTCCA 896  
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281  
DB 897 GCGCAATCTCAGGACCGGAATTCAGGCCGTTTCAACAAGAGGAGGACCTTGGC 956  
QY 282 TyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
DB 957 TACTATGATCTGTGACTTCTCCGCGAGGCCACAGTCCATAGAACCTCTCGCCAGCAG 1016  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValIleLysSerPheLys 320  
DB 1017 GTCCCTATGCCCACCAAGGGCAACAGTGGGTAGGATAGCAGCAGGAAGCGTCAAA 1076  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340  
DB 1077 AGCAAGGTGCAGTACCTGAAGGATAGGAGCTGGCAGCGCCATGATATGGGCTTGGAC 1136  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGly---ArgTyrProLeuIleGlnThr 359  
DB 1137 CTGGATGACTTCCAGGGCTCTCTTCTGCGGCCAGGATCTGCGCTTCCCTCTACCAATGCC 1196  
QY 360 LeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro 379  
DB 1197 ATCAAGGATGCACTCGCTGCAACG---TAGCCCTCTGTTCTGCACACAGCAGCGGGCCA 1253  
QY 380 LysProGlyGlnPro-----SerGluProGluHisGlyProSer-----Pro 393  
DB 1254 AGGATGCCCGCGTCTGCTGCTGGCGGAGAGCTGATCACCTGCCCTGCTGAGTCCCA 1313  
QY 394 Gly 394  
DB 1314 GGC 1316

Search completed: June 30, 2003, 01:11.43

Job time : 89.5244 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: June 29, 2003, 23:54:39 ; Search time 162.267 Seconds  
(without alignments)  
4070.941 Million cell updates/sec

Title: US-10-004-219B-10  
Perfect score: 2398  
Sequence: 1 AKLVCFYTNQAYRQGEARF.....QQSCPTGLVFNCKCCTWN 445

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 10 summaries

Command line parameters: -DEV-xlp  
-MODEL=frame.p2n.model -runat\_24062003\_160230\_8285/app.query.fasta\_1.1877  
-Q/cgn2\_1/USFTO.spool/US10004219/runat\_24062003\_160230\_8285/app.query.fasta\_1.1877  
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10004219 @CNG 1.1 123 @runat\_24062003\_160230\_8285  
-NCPUS=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOC=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FCAPOP=6 -FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1323	55.2	1625	9 US-10-004-219B-2
2	1301.5	54.3	1525	9 US-10-004-219B-3
3	1051	43.8	1925	9 US-10-097-340-44
4	1025	42.7	1681	10 US-09-215-077A-4

5	1025	42.7	1681	10	US-09-262-213A-4	Sequence 4, Appli
6	996	41.5	1391	10	US-09-822-830A-402	Sequence 402, App
7	951	39.7	1474	10	US-09-765-231A-8	Sequence 8, Appli
8	762	31.8	1470	9	US-10-218-743-40	Sequence 40, Appl
9	762	31.8	1470	9	US-10-218-743-42	Sequence 42, Appl
10	762	31.8	1527	9	US-10-218-743-37	Sequence 37, Appl
11	762	31.8	1527	9	US-10-218-743-39	Sequence 39, Appl
12	762	31.8	1621	9	US-10-218-743-34	Sequence 34, Appl
13	762	31.8	1621	9	US-10-218-743-36	Sequence 36, Appl
14	738	30.8	1608	9	US-10-218-743-20	Sequence 20, Appl
15	738	30.8	1608	9	US-10-218-743-22	Sequence 22, Appl
16	738	30.8	1665	9	US-10-218-743-17	Sequence 17, Appl
17	738	30.8	1665	9	US-10-218-743-19	Sequence 19, Appl
18	738	30.8	1752	9	US-10-218-743-14	Sequence 14, Appl
19	738	30.8	1752	9	US-10-218-743-16	Sequence 16, Appl
20	536	22.4	1140	9	US-09-938-842A-1684	Sequence 1684, Ap
21	524.5	21.9	1038	9	US-10-202-436A-14	Sequence 14, Appl
22	521.5	21.7	1797	10	US-09-974-300-646	Sequence 646, App
23	499	20.8	449	10	US-09-960-352-678	Sequence 678, App
24	487	20.3	1271	9	US-09-923-844B-1	Sequence 1, Appli
25	411	17.1	427	10	US-09-960-352-3315	Sequence 3315, Ap
26	410	17.1	438	10	US-09-960-352-1882	Sequence 1882, Ap
27	409	17.1	422	10	US-09-960-352-5957	Sequence 5957, Ap
28	409	17.1	423	10	US-09-960-352-12100	Sequence 12100, A
29	409	17.1	430	10	US-09-960-352-1189	Sequence 1189, Ap
30	404	16.8	426	10	US-09-960-352-4979	Sequence 4979, Ap
31	403	16.8	419	10	US-09-960-352-10172	Sequence 10172, A
32	402	16.8	422	10	US-09-960-352-10642	Sequence 10642, A
33	401	16.7	427	10	US-09-960-352-4978	Sequence 4978, Ap
34	399	16.6	419	10	US-09-960-352-4053	Sequence 4053, Ap
35	395	16.5	410	10	US-09-960-352-8317	Sequence 8317, Ap
36	394	16.4	429	10	US-09-960-352-14230	Sequence 14230, A
37	392	16.3	410	10	US-09-960-352-10161	Sequence 10161, A
38	390	16.3	440	10	US-09-960-352-4058	Sequence 4058, Ap
39	386	16.1	402	10	US-09-960-352-11601	Sequence 11601, A
40	384	16.0	424	10	US-09-960-352-3057	Sequence 3057, Ap
41	380	15.8	410	10	US-09-960-352-5496	Sequence 5496, Ap
42	378	15.8	437	9	US-09-918-995-35463	Sequence 35463, A
43	377	15.7	414	10	US-09-960-352-256	Sequence 256, App
44	374	15.6	399	10	US-09-960-352-265	Sequence 265, App
45	373	15.6	410	10	US-09-960-352-4536	Sequence 4536, Ap

ALIGNMENTS

RESULT 1  
US-10-004-219B-2  
; Sequence 2, Application US/10004219B  
; Publication No. US20030087414A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SEQ ID NO 2  
; LENGTH: 1625  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human  
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid  
; OTHER INFORMATION: sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase  
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence  
; FEATURE:



US-10-004-219B-10 (1-445) x US-10-004-219B-3 (1-1525)

Qy	3	LeuValCysTyrThrPheThrAsnTppAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro	22
Db	70	CTGATATGCTATTTCACCAACTGGGCCAGTATGCGCCAGGTCTGGGGAGCTTCAAGCCT	129
Qy	23	LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn	42
Db	130	GATGACATTAAACCCCTGGCTGTGTACTCACCTCATCTATGCTTTTGTCTGGATGCAGAAC	189
Qy	43	HisGlnLeuSerThrThrGluTrrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu	62
Db	190	AATGAGATCACCAACCATAGAAATGAATGATGTACTCTCTATAAAGCTTTCAATGACTTG	249
Qy	63	LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrrpAsnPheGlyThr	82
Db	250	AAAAACAGGAAACAGCAAACTGAATAACCTCTCTGGCAATTGGAGCTGGAACTTTGGAACT	309
Qy	83	GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla	102
Db	310	GCTCCTTTCACTACCATGGTTTCCACTTCTCAGAAACGCCAGACCTTCATTACTCAGTC	369
Qy	103	IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrrpGluTrrpGly	122
Db	370	ATCAAAATTTCTGGCTCAGTATGGTTTGTATGGACCTGGACCTGGACCTGGGAATACCCAGGC	439
Qy	123	SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla	142
Db	430	TCACGTGGGAGCCCTCTCCAGGACAGCATCTCTTCACTGCTCTGGTGAAGGAATGCGT	489
Qy	143	AsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAlaAla	162
Db	490	GAAGCTTTTTCAGCAGGAGGCTATTTCAGAGCAACAGAGCCAGACATGATGTGTACTGTGCT	549
Qy	163	ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn	182
Db	550	GTAGCTGGTGGATTTTCAACATCCAGCTGGCTGTATGAGATCCTCGAACCTTTCTAAGTAC	609
Qy	183	LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrrpGluLysValThrGly	202
Db	610	CTGGATTTTCATCCATGTGCATGACATATGACCTCCATGGCTCTCTGGGAGGCTACACTGGG	669
Qy	203	HisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeuAsnVal	222
Db	670	GAGAATAGTCCTCTTTACAAATACCCTTACTGAGACTGGTAGCAATGCTCACTCAATGTG	729
Qy	223	AspAlaAlaValGlnGlnTrrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGly	242
Db	730	GATTATGTCATGAACATTATGGAGAAACAATGGAGCCCCAGCTGAGAGAGCTCATTTGTGGA	789
Qy	243	MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla	262
Db	790	TTCCAGAGTATGGACACACCTTCATCCTCAGAAACCCCTCTGATAATGAAATTTGGTGCC	849
Qy	263	ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAlaTyr	282
Db	850	CCTACCTCTGTGTATGGCCCTCTGGCGCCCTATACACAGAGCGTGGGTCTCTGGGCCCTAC	909
Qy	283	TyrGluValCysSerTrrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys	300
Db	910	TATGAGATTTGCACCTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCCCAAGAA	969
Qy	301	ValProTrrpIlePheArgAspAsnGlnTrrpValGlyPheAspAspValGluSerPheLys	320
Db	970	GTGCCCTATGCTATAAGAGCCCAACAGTGGCTTGGCTATGACATATATCAAGAGCTTCAGT	1029
Qy	321	ThrLysValSerTrrpLeuLysGlnLysGlyLeuGlyAlaMetValTrrpAlaLeuAsp	340
Db	1030	GTTNAGGCTCAGTGGCTTAAGCAGAAACAATTTTGGAGGTGCCATGATCTCTGGCCATTTGAC	1089
Qy	341	LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu	360
Db	1090	CTTGATGACTTCACTGGCTCTTTCTGTGATCAGGCAAAATTTCTCTGACTTCTACTTTG	1149

Qy	361	ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys	380
		:::      :::	
Db	1150	AACAAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCTCACGTG	1197
		:::      :::	
Qy	381	ProGlnProSerGluPro	388
Db	1198	-----CCITCCAGCGCAGTGACTACTCTCCAGGAAGTGGGAGTGGGGTGGGAAC	1248
Qy	389	HisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro	408
Db	1249	TCCGGAGGAAGCTCTCGAGGACAGTGGATTCTGTGCGCAACAAAGCAGATGCGCTCTACCC	1308
Qy	409	AsnProArgGluAspSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSer	428
Db	1309	GTGGCAGATGACAGAATGCTTTTGGCAGTGCATCATGGAATCATACACAGCAGCAT	1368
Qy	429	CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp	444
Db	1369	TGTCAAGCAGGGCTGTGTTTTGATACACAGCTGAATTGCTGCAACTGG	1416

### RESULT 3

US-10-097-340-44  
; Sequence 44, Application US/10097340  
; Publication No. US20030087250A1

## GENERAL INFORMATION:

GENERATED BY: C:\COURT\CLERK	APPLICANT: JOHN MONAHAN	APPLICANT: MANJULA GANNANVARAPU
	APPLICANT: SEBASTIAN HOERSCH	APPLICANT: SEBASTIAN HOERSCH
	APPLICANT: SHUBHANGI KAMATKAR	APPLICANT: SHUBHANGI KAMATKAR
	APPLICANT: STEVE G. KOVATS	APPLICANT: STEVE G. KOVATS
	APPLICANT: RACHEL E. MEYERS	APPLICANT: RACHEL E. MEYERS
	APPLICANT: MICHAEL MORRISSEY	APPLICANT: MICHAEL MORRISSEY
	APPLICANT: PETER OLANDT	APPLICANT: PETER OLANDT
	APPLICANT: AMI SEN	APPLICANT: AMI SEN
	APPLICANT: PETER VEIBY	APPLICANT: PETER VEIBY
	APPLICANT: GORDON B. MILLS	APPLICANT: GORDON B. MILLS
	APPLICANT: ROBERT C. EAST, JR.	APPLICANT: ROBERT C. EAST, JR.
	APPLICANT: KAREN LU	APPLICANT: KAREN LU
	APPLICANT: ROSEMARIE SCHMANDT	APPLICANT: ROSEMARIE SCHMANDT
	APPLICANT: XUMEL ZHAO	APPLICANT: XUMEL ZHAO
	APPLICANT: KAREN GLATT	APPLICANT: KAREN GLATT

TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
 TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
 FILE REFERENCE: MRI-030

```

, FILE REFERENCE: MRI-030
, CURRENT APPLICATION NUMBER: US/10/097,340
, CURRENT FILING DATE: 2002-03-16
, PRIOR APPLICATION NUMBER: 60/276,025
, PRIOR FILING DATE: 2001-03-14
, PRIOR APPLICATION NUMBER: 60/325,149
, PRIOR FILING DATE: 2001-09-26
, PRIOR APPLICATION NUMBER: 60/276,026
, PRIOR FILING DATE: 2001-03-14
, PRIOR APPLICATION NUMBER: 60/324,967
, PRIOR FILING DATE: 2001/09/26
, PRIOR APPLICATION NUMBER: 60/311,732
, PRIOR FILING DATE: 2001-08-10
, PRIOR APPLICATION NUMBER: 60/325,102
, PRIOR FILING DATE: 2001-09-26
, PRIOR APPLICATION NUMBER: 60/323,580
, PRIOR FILING DATE: 2001-09-19
, NUMBER OF SEQ ID NOS: 363

```

NUMBER OF SEQ ID NOS: 303  
: SOFTWARE: FastSEO for Windows Version 4.0

; SOFTWARE: 1  
; SEQ ID NO: 4

REV ID NO 44  
: LENGTH: 1925

LENGTH: 1  
TYPE: DNA

TYPE: DNA  
ORGANISM: Homo sapiens

; ORGANISM: HO  
 10 007 340 41

Alignment Scores:		
Pred. No.:	2.12e-105	1925
Score:	1051.00	207
Percent Similarity:	68.56%	70
Conservative:		

Best Local Similarity: 51.24% Mismatches: 107  
Query Match: 43.83% Indels: 20  
DB: 9 Gaps: 10

US-10-004-219B-10 (1-445) x US-10-097-340-44 (1-1925)

```
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 193 AAACCTGGTCTGCTACTACACAGCTGGTCCAGTACCGGAGGCGATGGAGCTGCTTC 252
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41
DB 253 CCAGATGCCCTTCCAGCGCTCTCTCTGATCCACATCATCTACAGCTTTGCCAATAAAGC 312
QY 42 AsnHisGlnLeuSerThrThrGluTyrAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 313 AACGATCATCATGACACCTGGGAGTGGATGATGTGACGCTCTACGGCATGCTCAACACA 372
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTrpAsnPheGly 81
DB 373 CTCACGAACACAGGAACCCCAACCTGGAAGCTCTCTGTCTGCGGAGGATGGAACCTTTGG 432
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 433 TCTCAAGATTTTCCAGATAGCTCCCAACACACAGAGTGGCGGACTTTTCATCAAGTCA 492
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 493 GTACCGCCATCTCTGCGACCCATGCTTTGATGGCTGGACCTTCCCTGGCTTACCTT 552
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 553 GGACCGGAGA-----GACAAACAGCATTTTACCACCTTATCAAGGAATG 597
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 598 AAGCGCGAATTTATNAAGAAAGCCAG---CCAGGGAAGAAAGCAGCTCTGCTCAGCGCA 654
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 655 GCATCTCTGCGGGGAGGTCCACCTGACAGCATGTATGACATTCACCAAGATATCCCAA 714
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 715 CACCTGGATTTCATTAGCATGACCTACGATTTTCATGGAGCTGGCGTGGACCAACA 774
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeuAsn 221
DB 775 GGCCATCAGATCCCTCTGTCGAGTCCAGGAGGATGCAAGTCCCTGACAGATTCAGCAAC 834
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
DB 835 ACTGATATGCTGTGGGTACATGTTGAGGCTGGGGCTCTCGCCAGTAACTGGTGATG 894
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
DB 895 GGATCCCAACCTTCGGAGGAGCTTCACTTGCT---TCTCTGAGCATGCTGTGTGA 951
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281
DB 952 GCGCCCAATCTCAGGACCGGGAATTCAGCGCCGCTTCACCAAGAGGACGCGGACCTTTGC 1011
QY 282 TyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 1012 TACTATGATCTGTGACTCTCTCCGGGAGGCCACAGTCCATAGAACCTTCGCGCAGAG 1071
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
DB 1072 GTCCCCCTATGCCACCAAGCAACCAAGTGGTAGGATACGACGACGAGGAAGCGTCAAA 1131
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1132 AGCAAGGTGCAGTACCTGAAGGATAGGCGACTGGCAGGCGCCATGTATGGGCCCTTGGAC 1191
```

```
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGly---ArgTyrProLeuIleGlnThr 359
DB 1192 CTGGATGACTTCCAGGGCTCCTTCTGCGGCGAGGATGCGGCTTCCCTCTCACAATGCC 1251
QY 360 LeuArgGlnGluLeu-----SerLeuProTyrLeuProSerGlyThrPro 374
DB 1252 ATCAAGGATGCACTCGCTGCAACGTAAGCTCTGTCTGTGACACACAGCAGCGGG---CCA 1308
QY 375 GluLeuGluValProLysProGlyGlnProSer-----GluProGluHisGlyProSer 392
DB 1309 AGGATGCCCCGTCCC---CCTGTGCTCAGCTGGCGGAGCCTGATCACCTGCCCTGCG 1365
QY 393 -----ProGly 394
DB 1366 TGAGTCCCCAGGC 1377

RESULT 4
US-09-215-077A-4
; Sequence 4, Application US/09215077A
; Patent No. US20020031793A1
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; APPLICANT: JOHANSEN, JULIA S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; FILE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES
; FILE REFERENCE: 407T-895411US
; CURRENT APPLICATION NUMBER: US/09/215,077A
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/581,527
; PRIOR FILING DATE: 1996-04-17
; PRIOR APPLICATION NUMBER: 08/089,989
; PRIOR FILING DATE: 1993-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cDNA nucleotide
; OTHER INFORMATION: sequence for the coding region of the gene for
; OTHER INFORMATION: YKL-40.
US-09-215-077A-4

Alignment Scores:
Pred. No.: 1,25e-102 Length: 1681
Score: 1025.00 Matches: 202
Percent Similarity: 67.08% Conservative: 67
Best Local Similarity: 50.37% Mismatches: 116
Query Match: 42.74% Indels: 16
DB: 10 Gaps: 8

US-10-004-219B-10 (1-445) x US-09-215-077A-4 (1-1681)
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 138 AAACCTGGTCTGCTACTACACAGCTGGTCCAGTACCGGAGGCGATGGAGCTGCTTC 197
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41
DB 198 CCAGATGCCCTTTCAGCGCTTCTGTGTATCCCAACATCATCTACAGCTTTGCCAATAAAGC 257
QY 42 AsnHisGlnLeuSerThrThrGluTyrAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 258 AACGATCATCATGACACCTGGGAGTGGATGATGTGACGCTCTACGGCATGCTCAACACA 317
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTrpAsnPheGly 81
DB 318 CTCACAAACACAGGAACCCCAACCTGGAAGCTCTCTGTCTGTCGGAGGATGGAACCTTTGG 377
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 1192 CTGGATGACTTCCAGGGCTCCTTCTGCGGCGAGGATGCGGCTTCCCTCTCACAATGCC 1251
```

378 TCTCAAGATTTCACAGTAGCTCCCAACCCAGAGTGGCGGACTTTCATCAAGTCA 437  
QY 102 AlalileAArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121  
Db 438 GTACCGCCATTCTCGGACCCATGGCTTTGATGGCGTGACCTTGCTGGCTTACCTC 497  
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrLeuValGlnAspLeu 141  
Db 498 GGACGCGAGA-----GACAAACACCATTTTACCACCTAATCAAGAAATG 542  
QY 142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
Db 543 AAGGCCGAATTTATTAAGGAAGCCAG---CCAGGGAAGAAAGCAGCTCTGCTCAGCGCA 599  
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
Db 600 GCACCTGTCTGGGGAAGGTCACTACCATTCACAGCAGCTATGACATTGCCAGATATCCCAA 659  
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201  
Db 660 CACCTGGATTTCATTAGCATCATGACCTACGATTTTATGGCGCTGGCGTGGGACCA 719  
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsn 221  
Db 720 GGCCATCAGTCCCTCAGCGCGAGTCCAGGAGGATGCAAGTCTGACAGATTCAGCAAC 779  
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241  
Db 780 ACTGATATGCTGGGTGACATGTTGAGCTGGGGCTCTGCCAGTAAAGCTGGTGA 839  
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGly 261  
Db 840 GGCATCCACCTTCGGGAGGAGCTTCACTCTGGCT---TCTTCTGAGACTGGTGTCCA 896  
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281  
Db 897 GCGCAATCTCAGGACCGGGAATTCAGCGCGGTTTCAACAGGAGCGAGGACCTTGCC 956  
QY 282 TyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 957 TACTATGATGCTGTGACTTCTTCGCGGAGGACAGTCCATAGAACCTCGCGGACGAG 1016  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
Db 1017 GTCCCTATGCCACCAAGGCAACCACTGGGTAGGATACGACGACGAGAAAGCTCAA 1076  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
Db 1077 AGCAAGGTGCACTACTGTAAGGATAGGCGAGTGGCAGGCGCCATGATGGGCCCTGGAC 1136  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGly---ArgTyrProLeuIleGlnThr 359  
Db 1137 CTGGATGACTTCAGGGCTCTCTCGCGCCAGGATCTCGCTTCTCTCACCATGCCC 1196  
QY 360 LeuArgGlnGluSerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro 379  
Db 1197 ATCAAGGATGCACTCGCTCAAG---TAGCCCTCTGTTCTGCACACACGCGGGGCCA 1253  
QY 380 LysProGlyGlnPro-----SerGluProGluHisGlyProSer-----Pro 393  
Db 1254 AGGATGCCCGCTCCCGCTGCTGGCGGGAGCCTGATCACCTGCTGTGAGTCCCA 1313  
QY 394 Gly 394  
Db 1314 GGC 1316

## RESULT 5

US-09-262-213A-4  
; Sequence 4, Application US/09262213A  
; Patent No. US20020090658A1  
; GENERAL INFORMATION:  
; APPLICANT: PRICE, PAUL  
; APPLICANT: JOHANSEN, JULIA

; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF  
; TITLE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES  
; FILE REFERENCE: 407T-895412US  
; CURRENT APPLICATION NUMBER: US/09/262,213A  
; CURRENT FILING DATE: 1999-03-04  
; PRIOR APPLICATION NUMBER: US 08/089,989  
; PRIOR FILING DATE: 1993-07-09  
; PRIOR APPLICATION NUMBER: PCT/US94/07754  
; PRIOR FILING DATE: 1994-07-08  
; PRIOR APPLICATION NUMBER: US 08/581,527  
; PRIOR FILING DATE: 1996-04-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 4  
; LENGTH: 1681  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-262-213A-4

Alignment Scores:  
Pred. No.: 1,25e-102 Length: 1681  
Score: 1025.00 Matches: 202  
Percent Similarity: 67.08% Conservative: 67  
Best Local Similarity: 50.37% Mismatches: 116  
Query Match: 42.74% Indels: 16  
DB: 10 Gaps: 8

US-10-004-219B-10 (1-445) x US-09-262-213A-4 (1-1681)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
Db 138 AAACCTGGTCTGTACTACACAGCTGGTCCAGTACCGGGAAGCGATGGGAGCTGCTTC 197  
QY 22 ProLysAspLeuAspProSerLysCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
Db 198 CCAGATGCCCTTGACCGCTTCTGTGTATCCCACTATCATCAGCTTTGCCAATAAAGC 257  
QY 42 AsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
Db 258 AACGATCATCATCGACACCTGGGAGTGGGAATGATGTGCGCTCTACGGCATGCTCAACACA 317  
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
Db 318 CTCACCAACACAGAACCCCAACCTCGAAGACTCTCTTGTCTGCGGAGGATGGAACCTTGGG 377  
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
Db 378 TCTCAAGATTTTCCAGATAGCTTCCAAACCCACAGAGTGGCGGACTTTCATCAAGTCA 437  
QY 102 AlalileAArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121  
Db 438 GTACCGCCATTCTCGGACCCATGGCTTTGATGGCGTGACCTTGCTGGCTTACCTC 497  
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrLeuValGlnAspLeu 141  
Db 498 GGACGCGAGA-----GACAAACACCATTTTACCACCTAATCAAGAAATG 542  
QY 142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161  
Db 543 AAGGCCGAATTTATTAAGGAAGCCAG---CCAGGGAAGAAAGCAGCTCTGCTCAGCGCA 599  
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
Db 600 GCACCTGTCTGGGGAAGGTCACTACCATTCACAGCAGCTATGACATTGCCAGATATCCCAA 659  
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201  
Db 660 CACCTGGATTTCATTAGCATCATGACCTACGATTTTATGGCGCTGGCGTGGGACCA 719  
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsn 221  
Db 720 GGCCATCAGTCCCTCAGCGAGGTCCAGGAGGATGCAAGTCTGCTGACAGATTCAGCAAC 779

QY 222 ValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241  
Db 780 ACTGACTATGCTGGGGTACATGTTGAGGCTGGGGCTCTCCAGTAGCTGGTATG 839  
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261  
Db 840 GGATCCCACTTCGGGAGGAGCTTCACTCTGGCT---TCTTCTGAGACTGGTGTCCA 896  
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281  
Db 897 GCCTCAATCTCAGACCGGGAATTCAGGCCGGTTCACCAAGAGGAGGAGCCCTTGCC 956  
QY 282 TyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 957 TACTATGAGACTGTGACTTCTCCCGGAGGCCACAGTCCATAGAACCTCTCGGCCAGCAG 1016  
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320  
Db 1017 GTCCCTATGTCACCAAGGCCAACCAAGTGGGTAGGATACGACACGAGGAAGCGTCAA 1076  
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340  
Db 1077 AGCAAGTGCAGTACTTGAAGGATAGCAGCTGGCAGGCCCATGGTATGGGCCCTGGAC 1136  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGly---ArgTyrProLeuIleGlnThr 359  
Db 1137 CTGGATGACTTCCAGGCTCTTCTGGGCCAGGATCGGCTTCCCTCTCACCAATGCC 1196  
QY 360 LeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro 379  
Db 1197 ATCAAGATGCACTCGCTGCAACG---TAGCCCTCTGTCTGCACACAGCAGCGGGGCCA 1253  
QY 380 LysProGlyGlnPro-----SerGluProGluHisGlyProSer-----Pro 393  
Db 1254 AGATGCCCGCTCCCGCTGGCTGGCTGGCGGGAGCCCTGATCACCTGCCCTGCTGAGTCCCA 1313  
QY 394 Gly 394  
Db 1314 GGC 1316

## RESULT 6

US-09-822-830A-402  
; Sequence 402, Application US/09822830A  
; Patent No. US20020142952A1  
; GENERAL INFORMATION:  
; APPLICANT: Genetics Institute, Inc.  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulukota, Kamalakar  
; APPLICANT: Graham, James R.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6402  
; CURRENT APPLICATION NUMBER: US/09/822,830A  
; CURRENT FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/195,604  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 631  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 402  
; LENGTH: 1391  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1353  
; OTHER INFORMATION: n=a,c,g, or t

## US-09-822-830A-402

Alignment Scores:

Pred. No.: 1,45e-99 Length: 1391  
Score: 996.00 Matches: 195  
Percent Similarity: 63.51% Conservative: 80  
Best Local Similarity: 45.03% Mismatches: 126  
Query Match: 41.53% Indels: 32  
DB: 10 Gaps: 6

US-10-004-219B-10 (1-445) x US-09-822-830A-402 (1-1391)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
Db 79 AAACACTGGTGTGGAGCTTTACCAACTGGTCCCGAGGACCGGACGAAACCAAGAAATTCACC 138  
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
Db 139 CTTGAGATATATGACCCCTTCTCTATGCTCTCATCTCATCTATTCATTCGCCAGCATCGAA 198  
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61  
Db 199 AACAAACAGGTATTATCATCAAGGACCAAGAGTGAAGTGTCTTACACAGACCATCAACAGT 258  
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
Db 259 CTCAAACCAAGAAATCCCAACTGAAATTCCTTGTCCATTTGGAGGGTACTCTGTTGTGT 318  
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101  
Db 319 TCCAAAGGGTTCACCCCTATGTTGGATTCCTTCTACATCAGCTTGGAAATTCATTAATCTCC 378  
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121  
Db 379 ATAATCTCTGTTCTCAGGAAACCAATTAATTTGATGACTGGATGAAGCTGGATCTACCCA 438  
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141  
Db 439 GATCAGAA-----GAAACACTCAATTCCTGCTGCTGATTCATGAGTTA 483  
QY 142 AlaAsnAlaPheGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAla 161  
Db 484 GCAGAAGCTTTCAGAAGGACTTCAAAATCCACCAAGGAAGGCTTCTTGTAGTCGG 543  
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181  
Db 544 GCGGTATCTGCAGGAGGCAAAATGATGATACAGCTATCAAGTTTCAGAAACTGGCAAAA 603  
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLys----- 199  
Db 604 GATCTGGATTTTCATCACTCTGCTTGTACTTCCATGGGTCTTGGGAAAGGCCCTTT 663  
QY 200 ValThrGlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyAlaAlaAsp 219  
Db 664 ATCACTGGCCCAACACAGCCCTCTGAGCAAGGGGTGGCAGGACAGAGGCCCAAGCTCTTAC 723  
QY 220 LeuAsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 239  
Db 724 TACAATGTGGAATATGCTGTGGGGTACTGGATACATAAGGGAATCCATCAGAGAAGGTG 783  
QY 240 IleLeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArg 259  
Db 784 GTCATGGGATCCCCACATATGGGCACCTCTTCACTGGGCC---TCTGCAAGAAACCAACC 840  
QY 260 ValGlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMet 279  
Db 841 GTGGGGGCCCTGCTCTGGCCCTGGAGCTGTGGACCCCATCATCAGAGTCTTTCAGGCTTC 900  
QY 280 LeuAlaTyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAsp 298  
Db 901 CTGGCCTATATGAGATCTGCCAGTTCCTCTGAAAGGAGCCAGATCAGCAGGCTCCAGAT 960  
QY 299 GlnLysValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSer 318  
Db 961 CAGCAGGTTCCTCAGCAGTCCAGGGGGAGCCCGTGGGTGGGCTATGATGATGTGAAGAGT 1020



QY 319 PhetythrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrrpAla 338  
Db 1021 TTGGGGCCCAAGTTTCAGTTCTTAAAGAAATTAAACCTGGGGGGTGCCTGATTGGTCT 1080  
QY 339 LeuAspLeuAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGln 358  
Db 1081 TTTGACATGGTTGACTTCACTGGCAAACTCTGCAACCGGGGCCCTTCCCTCTTGTGCCAA 1140  
QY 359 ThrLeuArgGlnGluLeu----- 364  
Db 1141 GCAGTCAGAGAGAGCTTGGCTCCCTGTGAAGGATTAACCTACAGAGAGAGGCAAGAT 1200  
QY 365 -----SerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381  
Db 1201 GCCCTTGTGCTGGGGCTGCTCTCTCCAGGAATTCATGTGGGATTCCTCCCTTGCCA 1260  
QY 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAsp----- 396  
Db 1261 GGCCGGCCTTTGGATCTCTCTTCCAAGCCTTTCTTGACTTCTCTTAGATCATAGATTGG 1320  
QY 397 ---ThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro 408  
Db 1321 ACCTGGTTTGGTTTCTTCGACGCTGATGCTTNTTGGCC 1359  
RESULT 7  
US-09-765-231A-8  
; Sequence 8, Application US/09765231A  
; Patent No. US20020119452A1  
; GENERAL INFORMATION:  
; APPLICANT: Searle/Monsanto  
; APPLICANT: Phippard, Deborah  
; APPLICANT: Vasanthakamur, Geetha  
; APPLICANT: Dotson, Stanton  
; APPLICANT: Ma, Xiao-Jun  
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,  
; FILE OF INVENTION: vectors, and cells  
; FILE REFERENCE: SO-3221 PR  
; CURRENT APPLICATION NUMBER: US/09/765,231A  
; CURRENT FILING DATE: 2001-01-18  
; NUMBER OF SEQ ID NOS: 82  
; SEQ ID NO 8  
; LENGTH: 1474  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-765-231A-8  
Alignment Scores:  
Pred. No.: 1.37e-94 Length: 1474  
Score: 951.00 Matches: 196  
Percent Similarity: 66.19% Conservative: 80  
Best Local Similarity: 47.00% Mismatches: 108  
Query Match: 39.66% Indels: 33  
Gaps: 5  
US-10-004-219B-10 (1-445) x US-09-765-231A-8 (1-1474)  
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhele 21  
Db 140 AAACCTGGTTTGTACTTTTACCACTGGGTCCAGGACCGGCAAGCAACCGAAATTCAC 199  
QY 21 uPro-LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPhe-AlaGlyMet 40  
Db 200 CCTGAGGAATATGACCCCTTCTCTATGCTCTCATCTCATCTATTCAATTCGCCAGCATC 259  
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrrpAsnAspGluThrLeuTyrGlnGluPheAsn 60  
Db 260 GAAACACACAGAGTTATCATCAAGGACACAGAGTGAAGTGTGCTCTACACAGACCATCAAC 319  
QY 61 Gly-LeuLysLysMetAsnProLysLeuLysThrHisLeuAlaIleGlyGlyTrrpAsnPh 80  
Db 320 AGTTCTCAAAACCAAGAAATCCCAAAATTCCTTGTCTCCATGGAGGATCCTGTT 379  
QY 80 eGlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAs 100

Db 380 TGGTTCCAAAGGGTTCCACCCTATGGTGGATTCTTACATCAGGCTTGAATTCAATAA 439  
QY 100 nSerAlaIleArgPheLeuArgLysTyrSerPheAsnGlyLeuAspLeuAspTrrpGluTy 120  
Db 440 CTCATATCTCTGTTCTTGAGGAACCACTAACTTTGATGGACTGGATGAAGCTGGATCTA 499  
QY 120 rProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAs 140  
Db 500 CCCAGATCAGAAA-----GAAACACTCATTTCTACTGTGCTGATTATCA 544  
QY 140 pLeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSe 160  
Db 545 GTTAGCAGAACCTTTCAAGAGGACTTCACAAAATCCACCAAGGAAGGCTTCTCTTGAC 604  
QY 160 rAla-AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrClnValAspLysLea 180  
Db 605 TGCGGGGGGTATCTGCAGGGAGGCAATATGATGATAACAGCTATCAAGTTGAGAACTGG 664  
QY 180 laGlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrrpGluLys- 199  
Db 665 CAAAGATCTGGATTTCATCAACCTCTGCTCTTGTACTTCCATGGGTCTTGGGAAAGC 724  
QY 200 -----ValThrGlyHisAsnSerPro-LeuTyrLysArgGlnGluGluSerGlyAlaAla 217  
Db 725 CCTTATCACTGGCCACAAACAGCCCTGCTGAGCAAGGGTGGCAGGACAGAGGGCCAAGC 784  
QY 218 AlaSerLeuAsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSer 237  
Db 785 TCCTACTACAATGTGGAATATGCTGTGGGTACTGTGATACATATAAGGGAATGCCATCAG 844  
QY 238 LysLeuIleLeuGlyMetProThrTyr-GlyArgSerPheThrLeuAlaSerSerSerAs 257  
Db 845 AAGTGTGTCATGGGATCCCCACATATGCGGCACCTCTTCACTGGCC---TCTGCGAGA 901  
QY 257 pThrArgValGlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGl 277  
Db 902 AACCACTGGGGGCCCTGCTGCGCCCTGGAGCTGCTGAGCCCATCACAGAGCTTTC 961  
QY 277 yGlyMetLeuAlaTyrTrrpGluValCysSerTrp---LysGlyAlaThrLysGlnArgIl 296  
Db 962 AGGCTTCTGCTTATATGATGATCTGCCAGTCTCTGAAAGGAGGCAAGATCACCGGCT 1021  
QY 296 eGlnAspGlnLysValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspVa 316  
Db 1022 CCAGGATCAGCAGGTTCCCTACGAGTCAAGGGNACCAGTGGGTGGGTATGATGATGT 1081  
QY 316 lGluSerPheLysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetVa 336  
Db 1082 GAAGAGTATGGAGACCAAGTTTCAGTTCTTAAAGAAATTTAAACCTGGGAGGAGCCATGAT 1141  
QY 336 lTrrpAlaLeuAspLeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProle 356  
Db 1142 CTGGTCTATTACATGGATGATCTTCACTGGCAAAATCTCTGCAACCCAGGGCCCTTACCTCT 1201  
QY 356 uIleGlnThrLeuArgGlnGluLeu----- 364  
Db 1202 TGTCCACAGCTCAAGAGAAGCTTGGCTCCCTGTGAAGGATTAACCTTACAGAGAGCAG 1261  
QY 365 -----SerLeuProTyrLeuProSerGlyThrProGluLeuGluValPr 379  
Db 1262 GCAAGATCACCTTGTGCTGGGGCTGCTCTCTCCAGGAATTCATGTGGGATTCCT 1321  
QY 379 oLysProGlyGlnProSerGluProGluHisGlyPro 391  
Db 1322 CTGCGAGCGGGCCTTTGGATCTCTCTTCCAAGCCT 1358  
RESULT 8  
US-10-219-743-40  
; Sequence 40, Application US/10218743  
; Publication No. US2003009679A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine A.

QY 170 ValAspAlaGlyTyrGluValAspIleLeuAlaGlnAsnLeuAspPheValIleLeuMet 189  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 556 ATTACGTAGCTATTAGCTCAAGAATTAACAATAATTCGATTGGATGAATGCATG 615  
 QY 190 AlaTyAspPheHisGlySerTrpGluLysValThrGlyHisAsnSerProLeuTyrLys 209  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 616 ACTTATGATTACCATCGCGGATGGGAATAATTTTCGCCCAATAAGCTCCGTTGATAAA 675  
 QY 210 ArgGlnGluGluSerGlyAlaAlaSer---LeuAsnValAspAlaValGlnGln 228  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 676 CGACCGATGAACGGATGAATTCACACTTACTTCAATGTCACCATCAGCATAT 735  
 QY 229 TrpLeuGlnLysGlyThrProAlaSerLysLeuLeuLeuGlyMetProThrTyrGlyArg 248  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 736 TATTGAACAATGGCGCTACTCGAGACAACTTGTTATGGGTGTTCCATTCTATGTCGT 795  
 QY 249 SerPheThrLeuAlaSerSerSerAspThrArgValGlyAlaProAlaThrGlySerGly 268  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 796 GCTTGGAGCATGAAGATCGAAGCAAAAGTCAAATGGCGATCCGGCCAAAGCATGTCT 855  
 QY 269 ThrProGlyProPheThrLysGluGlyGlyMetLeuAlaTyrTrpGluValCysSerTrp 288  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 856 CTCCTCGTTTTATTACTGGTGGAGNAGTGTCTCTCATCATCGAATGTGTCAAGTTA 915  
 QY 289 LysGlyAlaThrLysGlnArgIleGln---AspGln-----LysValProTyrIlePhe 305  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 916 TTCCAGAAAAGAAATGGCATATTCAATACGATGAATATTCAATGCTCCATACGGATAT 975  
 QY 306 ArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThrLysValSerTyr 325  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 976 AATGATAAAATCTGGGTGGTTACGATGATCTGGCTAGTAGTATATCATGCAAGTTGGCCTT 1035  
 QY 326 LeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeuAspPheAla 345  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 1036 CTCAAGAATTCGGGCTCTCGCGTATGATGTGTCATGTGGAACACGATGATTTCAA 1095  
 QY 346 GlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArgGlnGluLeuSer 365  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 1096 GGT---CATTCGGGCGG--AAATATCCATTGTTGAACAAAGTTCACAATATGATCAAT 1149  
 QY 366 -----LeuPro 367  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 1150 GGATGATGAAGAAGCTCTTACGAATGTCTTTTGGGCCCAAGTACACCAACCAACCA 1209  
 QY 368 TyrLeuProSerGlyThrProGluLeuGluValProLysProGlyGlnProSerGluPro 387  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 1210 ACCACCGCTCAACTACT-----TCGACTACCACCAACCGCTACCACCGAT 1260  
 QY 388 GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLys-----402  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 1261 AGCACAAAGCAACCAACCAAAATACACTACGTATTTGATGGACATTTGATTAATGCTAT 1320  
 QY 403 AlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCys-----419  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 1321 AAACAAGGTTATCTTCCACACTCACTGATGTTTCATAAATATTTAGTTTGTGAATATTT 1380  
 QY 420 -----AlaAlaGlyArgLeuPheGlnSerCysProThrGlyLeuValPhe 435  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 1381 GCCACACCAACCGGTGGTGGGTGGGTGATGATGATGATGATGATGATGATGATGATG 1440  
 QY 436 SerAsnSerCysLysCysCys 442  
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 1441 CAGCAACATTAATAAATTTGT 1461

RESULT 9  
US-10-218-743-42/c  
; Sequence 42, Application US/10218743  
; Publication No. US20030096779A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine A.  
; APPLICANT: Hunter, Shirley Wu  
; APPLICANT: Weber, Eric R.

RESULT 9  
US-10-218-743-42/c  
; Sequence 42, Application US/10218743  
; Publication No. US20030096779A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine A.  
; APPLICANT: Hunter, Shirley Wu  
; APPLICANT: Weber, Eric R.



; PRIOR APPLICATION NUMBER: US/09/292,225  
 ; PRIOR FILING DATE: 1999-04-15  
 ; PRIOR APPLICATION NUMBER: 60/098,909  
 ; PRIOR FILING DATE: 1998-09-02  
 ; PRIOR APPLICATION NUMBER: 60/085,295  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/098,565  
 ; PRIOR FILING DATE: 1998-04-17  
 ; PRIOR APPLICATION NUMBER: 09/062,013  
 ; PRIOR FILING DATE: 1998-04-17  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 37  
 ; LENGTH: 1527  
 ; TYPE: DNA  
 ; ORGANISM: Dermatophagoides farinae  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(1527)  
 ; US-10-218-743-37

Alignment Scores:  
 Pred. No.: 8,1e-74 Length: 1527  
 Score: 762.00 Matches: 162  
 Percent Similarity: 52.36% Conservative: 93  
 Best Local Similarity: 33.26% Mismatches: 172  
 Query Match: 31.78% Indels: 60  
 DB: 9 Gaps: 14

US-10-004-219b-10 (1-445) x US-10-218-743-37 (1-1527)

QY 2 LysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
 DB 100 CGAATCGATGTTATGTTGGAAACATGTCGCTGTTATCATAAAGTTGATCCA---TACACA 156  
 QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41  
 DB 157 ATTGAAGATATTGATCCTTTCAAAATGTTACTCATTTGATGTTATGTTTGTCTAAATCGAT 216  
 QY 42 AsnHisGlnLeuSerThr-----ThrGluTyr 50  
 DB 217 GAATACAAATACACATTCATTCAGTTTGTATCCATTTCAAGATGATTAACATACTCATGG 276  
 QY 51 AsnAspGluThrLeuTyrGlnGluPheAsnGlyLeuLysLysMetAsnProLysLeuLys 70  
 DB 277 ---GAAAAACACCGGTATGAACGTTTCAACAACTTGAGATTGAAGATCCAGAATTGACC 333  
 QY 71 ThrLeuLeuAlaIleGlyThrAsnGlyThrGlnLysPheThrAspMetValAla 90  
 DB 334 ACCATGATTTCAATGGGTGTTGTTATGAAGGTTTCAGAAAAATATTTCGGATATGGCAGCC 393  
 QY 91 ThrAlaAsnAsnArgGlnThrPheValAsnSerAlaIleArgPheLeuArgLysTyrSer 110  
 DB 394 ATCCAAACATATCGTCAGCAATTTGTTCAATAGTTTGGACATTTTTCAGAAATAACAA 453  
 QY 111 PheAspGlyLeuAspLeuAspTrpGluTyrProGlySerGln---GlySerProAlaVal 129  
 DB 454 TTCGATGGCCCTAGATTGGATTGGGAATATCTCGGATCAGCGTTAGGCAATCTTAAATC 513  
 QY 130 AspGlyGluArgPheThrThrLeuValGlnAspLeuAlaAsnAlaPheGlnGlnGluAla 149  
 DB 514 GATAAACAACAACTATTAAACATAGTTAGAGAACTTAAAGAGGCAATTTGAACCTTTTCGGC 573  
 QY 150 GlnThrSerGlyLysGluArgLeuLeuSerAlaAlaValProAlaGlyGlnThrTyr 169  
 DB 574 -----TACTTTGTTGACCTGCCGAGTATCCACCCGGTAAGATAA 612  
 QY 170 ValAspAlaGlyTyrGluValAspLysIleAlaGlnAsnLeuAspPheValAsnLeuMet 189  
 DB 613 ATTGACGTAGCTTATGAGCTCAAGAAGATTGAACCAATTGTTGATTTGGATGATGTCATG 672  
 QY 190 AlaTyrAspPheHisGlySerTrpGluLysValThrGlyHisAsnSerProLeuTyrLys 209

DB 673 ACTTATGATTACCATGGCGGATGGGAAATGTTTTCGGCCATAATGCTCCGTTGTATAAA 732  
 QY 210 ArgGlnGluIleSerGlyAlaAlaSer---LeuAsnValAspAlaAlaValGlnGln 228  
 DB 733 CGACCCGATGAACCGATGCAATTCACACTTACTTCAATGTCACTACACCATGCACAT 792  
 QY 229 TrpLeuGlnLysGlyThrProAlaSerLysLeuLeuLeuGlyMetProThrTyrGlyArg 248  
 DB 793 TATTTGAACAATGGCGCTACTCGAGACAACTTGTATGGGTGTTCCTTCTATGTCGT 852  
 QY 249 SerPheThrLeuAlaSerSerSerAspThrArgValGlyAlaProAlaThrGlySerGly 268  
 DB 853 GCTTGGAGCATCGAAGATCGAAGCAAGTCAACTTGGCGATCCGCGCAAGAGCATGCT 912  
 QY 269 ThrProGlyProPheThrLysGlyGlyMetLeuAlaTyrTyrGluValCysSerTrp 288  
 DB 913 CCTCTGGTTTATTACTGCTGGAAGAGTCTCTCTCATACATCGAATTTGTGTCAGTTA 972  
 QY 289 LysGlyAlaThrLysGlnArgIleGln---AspGln-----LysValProTyrIlePhe 305  
 DB 973 TTCCAGAAAAGAAATGTCATATTCATACGATGAATATTAATGCTCCATACGATAT 1032  
 QY 306 ArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThrLysValSerTrp 325  
 DB 1033 AATGATAAAATCTGGTGGTTTACGATGCTGCTGATATATCATGCAAGTTGGCCTTT 1092  
 QY 326 LeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeuAspPheAla 345  
 DB 1093 CTCAAAGAAATTCGGCGCTCTCTGGCGTTATGATGTCATTCGTAACGATGATTTCAA 1152  
 QY 346 GlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArgGlnGluLeuSer 365  
 DB 1153 GGT---CATTCGGGACCG---AAATATCAATGTTGAACAAAGTTTCACAATATGATCAAT 1206  
 QY 366 -----LeuPro 367  
 DB 1207 GGTGATGAAAAGAACTTTCAGAAATGCTTTTGGGCCCAAGTACCAACCAACACACCA 1266  
 QY 368 TyrLeuProSerGlyThrProGluLeuGluValProLysProGlyGlnProSerGluPro 387  
 DB 1267 ACCACCCGTCACACTACT-----TCGACTACCACCAACCGCTTACCACCCAGT 1317  
 QY 388 GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLys----- 402  
 DB 1318 AGCACAAGCGAAACCAACCAAAATACACTACGATATATGATGGACATTTGATTAATGCTAT 1377  
 QY 403 AlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCys----- 419  
 DB 1378 AAACAAGGTTATCTTCCACATCCAACTGTTCAATAATATTTAGTTTGTGAATATATT 1437  
 QY 420 -----AlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPhe 435  
 DB 1438 GCCACACCAACCGTGGTGGGTGATGATGATGATGATGATGATGATGATGATGATGATG 1497  
 QY 436 SerAsnSerCysLysCysCys 442  
 DB 1498 CACGCAACATTAATAAATGTT 1518

# RESULT 11

US-10-218-743-39/c  
 ; Sequence 39, Application US/10218743  
 ; Publication NO. US20030096779A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McCall, Catherine A.  
 ; APPLICANT: Weber, Eric R.  
 ; APPLICANT: Hunter, Shirley Wu  
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
 ; TITLE OF INVENTION: AND USES THEREOF  
 ; FILE REFERENCE: AL-2-03  
 ; CURRENT APPLICATION NUMBER: US/10/218,743  
 ; CURRENT FILING DATE: 2002-08-13  
 ; PRIOR APPLICATION NUMBER: US/09/292,225  
 ; PRIOR FILING DATE: 1999-04-15



; PRIOR FILING DATE: 1998-04-17  
 ; PRIOR APPLICATION NUMBER: 09/062,013  
 ; PRIOR FILING DATE: 1998-04-17  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 34  
 ; LENGTH: 1621  
 ; TYPE: DNA  
 ; ORGANISM: Dermatophagoides farinae  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (14)..(1540)  
 ; US-10-218-743-34

Alignment Scores:  
 Pred. No.: 8,83e-74 Length: 1621  
 Score: 762.00 Matches: 162  
 Percent Similarity: 52.36% Conservative: 93  
 Best Local Similarity: 33.26% Mismatches: 172  
 Query Match: 31.78% Indels: 60  
 DB: 9 Gaps: 14

US-10-004-219b-10 (1-445) x US-10-218-743-34 (1-1621)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21  
 DB 113 CGAATCGTATGTTATGTCGAACATGTCGGTTTATCATAAAGTTGCATCA---TACACA 169  
 QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41  
 DB 170 ATTGAAGATATTGATCTTTCAAATGTACTCATTTGATGTTGTTTCTGCTAAATCGAT 229  
 QY 42 AsnHisGlnLeuSerThr-----ThrGluTrp 50  
 DB 230 GAATACAAATACACCATTCAGTTTGTGATTCATTTCAAGATGATCAACATCACTCATGG 289  
 QY 51 AsnAspGluThrLeuTyrGlnGluPheAsnGlyLeuLysLysMetAsnProLysLeuLys 70  
 DB 290 ---GAAACACCGGTATGAACCGTTTCAACACTTGAGATTGAAGAAATCCAGAAATTGACC 346  
 QY 71 ThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrGlnLysPheThrAspMetValAla 90  
 DB 347 ACCATGATTTCATTTGGTGGTGGTATGAAGGTTTCAGAAATATTTCCGATATGGCAGCC 406  
 QY 91 ThrAlaAsnAsnArgGlnThrPheValAsnSerAlaIleArgPheLeuArgLysTyrSer 110  
 DB 407 AATCCACATATCGTCAGCAATTTGTTCAATCACTTTGACCTTTTTCAGAAATACAA 466  
 QY 111 PheAspGlyLeuAspLeuAspTrpGluTyrProGlySerGln---GlySerProAlaVal 129  
 DB 467 TTCGATGGCTAGATTTCGATTTGGGAATATCTGGGATCACGGTTAGGCAATCTTAAATC 526  
 QY 130 AspLysGluArgPheThrThrLeuValGlnAspLeuAlaAsnAlaPheGlnGluAla 149  
 DB 527 GATAACAAAACTATTAACTAGTTAGAGAACTTAAAGAGCGATTCGAACCTTTCGCG 586  
 QY 150 GlnThrSerGlyLysGluArgLeuLeuSerAlaAlaValProAlaGlyThrThrTyr 169  
 DB 587 -----TACTTGTGACTGCGCGAGTATCACCCTCGTAAAGTAA 625  
 QY 170 ValAspAlaGlyTyrGluValAspLysIleAlaGlnAsnLeuAspPheValAsnLeuMet 189  
 DB 626 ATTGAGCTAGCTTATGAGCTCAAGAATTTGAACCAATTTGATGATGATGATGATG 685  
 QY 190 AlaTyrAspPheHisGlySerTrpGluLysValThrGlyHisAsnSerProLeuTyrLys 209  
 DB 686 ACTTATGATTACCATGGCGGATGGGAAATGTTTCGGCCATAAATGCTCCGTGTATATA 745  
 QY 210 ArgGlnGluSerGlyAlaAlaAlaSer---LeuAsnValAspAlaAlaValGlnGln 228  
 DB 746 CGACCCGATGAACCGATGAATGTCACATCTTACTTCAATGTCACTACACCATGCACTAT 805  
 QY 229 TrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGlyMetProThrTyrGlyArg 248

DB 806 TATTTGAACAATGGCGCTACTCGACACAACTTGTATGGGTGTTCCATTCTATGGTCT 865  
 QY 249 SerPheThrLeuAlaSerSerSerAspThrArgValGlyAlaProAlaThrGlySerGly 268  
 DB 866 GCTTGGAGCATCGAAGATCGAAGCAAGTCAAACTTGGGATCCGCGCAAGGCATGCT 925  
 QY 269 ThrProGlyProPheThrLysGluGlyMetLeuAlaTyrTyrGluValCysSerTrp 288  
 DB 926 CCTCTCTGTTTATTACTTGGTGAAGAAGTGTCTCTCATACATCGAATTTGTGAGTTA 985  
 QY 289 LysGlyAlaThrLysGlnArgIleGln---AspGln-----LysValProTyrIlePhe 305  
 DB 986 TTCCAGAAAGAAGATGGCATATTTCATACGATGAATATTACAATGCTCCATACCGATAT 1045  
 QY 306 ArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThrLysValSerTrp 325  
 DB 1046 AATGATAAAATCTGGTGGTACGATGATCTGGGTAGTATATCATGCAAGTGGCCTTT 1105  
 QY 326 LeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeuAspPheAla 345  
 DB 1106 CTCNAAGAATTGGGCGTCTCTGGCGTTATGATGTCATTGGAAACGATGATTTCAA 1165  
 QY 346 GlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArgGlnGluLeuSer 365  
 DB 1166 GGT---CATTCGGACCG---AAATATCCATTGTTTGAACAAAGTTTCACAATATGATCAAT 1219  
 QY 366 -----LeuPro 367  
 DB 1220 GGTGATGAAGAAGAACTTTCAGATGCTTTTGGGCCCAAGTACAAACACCAACACCA 1279  
 QY 368 TyrLeuProSerGlyThrProGluLeuGluValProLysProGlyGlnProSerGluPro 387  
 DB 1280 ACCACCCGCTCACTACT-----TCGACTACCACCAACGCTTACCACCCGAT 1330  
 QY 388 GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLys----- 402  
 DB 1331 AGCACAAGCGAAACACCAACAAATACACTAGTATATTGATGGACATTTGATTAATGCTAT 1390  
 QY 403 AlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCys----- 419  
 DB 1391 AAACAAGTATTCTTCCACATCCCACTGATGTTCTATAAATATTAGTTTGTGAATATT 1450  
 QY 420 -----AlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPhe 435  
 DB 1451 GCCACACCAACCGGTGGTGGTGTCACATATTGATTTGTCAAAAGGAAGTATAGTGG 1510  
 QY 436 SerAsnSerCysLysCysCys 442  
 DB 1511 CACGCAACATTAAATAATGT 1531

# RESULT 13

US-10-218-743-36/c  
 ; Sequence 36, Application US/10218743  
 ; Publication No. US20030096779A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McCall, Catherine A.  
 ; APPLICANT: Hunter, Shirley Wu  
 ; APPLICANT: Weber, Eric R.  
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
 ; FILE OF INVENTION: AND USES THEREOF  
 ; FILE REFERENCE: AL-2-C3  
 ; CURRENT APPLICATION NUMBER: US/10/218,743  
 ; CURRENT FILING DATE: 2002-08-13  
 ; PRIOR APPLICATION NUMBER: US/09/292,225  
 ; PRIOR FILING DATE: 1999-04-15  
 ; PRIOR APPLICATION NUMBER: 60/098,909  
 ; PRIOR FILING DATE: 1998-09-02  
 ; PRIOR APPLICATION NUMBER: 60/085,295  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/098,565  
 ; PRIOR FILING DATE: 1998-04-17  
 ; PRIOR APPLICATION NUMBER: 09/062,013







Db	573	CTGAAGAATAAGCGCTTTTGGTGTGTCATGGTTTGGTCATTGGAAATGATGATTTCAA	514
Qy	346	Gly-----	346
Db	513	GGTCACTCGGACCGAAAATCCATTGTTGAACAAAGTTCATAATATGATTAATGGCGAT	454
Qy	347	-----PheSerCys-----	353
Db	453	GAAGAAGACTCTTTTCGAATGCATTTTGGGTCCAAGTACAACAGCACCAACTCCAAACGACG	394
Qy	353	gTyProLeuIleGlnThrLeuArgGlnGluLeuSerLeuProTyLeuProSerGlyTh	373
Db	393	ACACCCACACACCCCGACTACAACGCCAACAACTCTCTCCACC-ACCCCGACACAACAC	335
Qy	373	rProGluLeuGluValPro---LysProGlyGlnProSerGluProGluHisGlyProse	392
Db	334	CCCTTCTCCACACCCCGACACAACACCCCTTCTCCACACACCGACACAACACTCCTTC	275
Qy	392	rProGlyGlnAspThr	397
Db	274	TCCACCCACACCAACA	259

Search completed: June 30, 2003, 03:58:16  
Job time : 186.267 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:10:28 ; Search time 1466.66 Seconds

(without alignments)  
4913.891 Million cell updates/sec

Title: US-10-004-219B-10

Perfect score: 2398

Sequence: 1 AKLVCFYFTNWAQYRQGEARF.....QQSCPTGLVFNCKCCTWN 445

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q/cgn2\_1/USPTO spoof/US10004219/runat\_24062003\_160228\_8173/app query.fasta\_1.1877  
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CGN\_1\_1\_4575 @runat\_24062003\_160228\_8173 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: 1: em\_estba: 2: em\_esthum: 3: em\_estin: 4: em\_estmu: 5: em\_estov: 6: em\_estpl: 7: em\_estro: 8: em\_htc: 9: gb\_est1: 10: gb\_est2: 11: gb\_hic: 12: gb\_est3: 13: gb\_est4: 14: gb\_est5: 15: em\_estfun: 16: em\_estom: 17: gb\_gss: 18: em\_gss\_hum: 19: em\_gss\_inv: 20: em\_gss\_pln: 21: em\_gss\_vrt: 22: em\_gss\_fun: 23: em\_gss\_mam: 24: em\_gss\_mus: 25: em\_gss\_other: 26: em\_gss\_pro: 27: em\_gss\_rod:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1471	61.3	1581	11	AK009012	AK009012 Mus muscu
2	1309.5	54.6	1533	11	AK008633	AK008633 Mus muscu
3	1304.5	54.4	1535	11	AK008757	AK008757 Mus muscu
4	1249.5	52.1	1534	11	AK008650	AK008650 Mus muscu
5	1207.5	50.4	1530	11	AK007573	AK007573 Mus muscu
6	1062	44.3	973	13	BI411922	BI411922 602966302
7	960.5	40.1	837	13	BI412518	BI412518 60290786
8	906	37.8	892	14	BQ231098	BQ231098 AGENCOURT
9	859	35.8	838	12	BG867815	BG867815 602786336
10	832	34.7	872	14	BQ231064	BQ231064 AGENCOURT
11	831	34.7	837	12	BG868741	BG868741 602787570
12	814	33.9	763	12	BG872647	BG872647 602793705
13	808	33.7	764	12	BG869304	BG869304 602788819
14	808	33.7	880	12	BG871468	BG871468 602790678
15	806	33.6	1098	13	BM458470	BM458470 AGENCOURT
16	804	33.5	833	12	BG871337	BG871337 602790509
17	802	33.4	894	12	BG869958	BG869958 602789826
18	798.5	33.3	939	14	BQ219632	BQ219632 AGENCOURT
19	798	33.3	886	12	BG870840	BG870840 602791891
20	796	33.2	847	12	BG866775	BG866775 602786605
21	794	33.1	855	12	BG872136	BG872136 602792974
22	793	33.1	948	13	BI664452	BI664452 602390028
23	791.5	33.0	842	12	BG866831	BG866831 602786666
24	790.5	33.0	848	12	BG869143	BG869143 602789020
25	790.5	33.0	976	14	BQ225554	BQ225554 AGENCOURT
26	787.5	32.8	822	12	BG870367	BG870367 602791360
27	787.5	32.8	836	12	BG869655	BG869655 602789716
28	783	32.7	838	12	BG870392	BG870392 602791392
29	781	32.6	987	9	AL547779	AL547779 AL547779
30	780.5	32.5	837	12	BG865578	BG865578 602783726
31	779.5	32.5	824	12	BG865859	BG865859 602788278
32	779.5	32.5	882	12	BG873488	BG873488 602791730
33	779.5	32.5	949	12	BG868019	BG868019 602788381
34	779	32.5	809	12	BG866417	BG866417 602785352
35	778.5	32.5	838	12	BG868956	BG868956 602784563
36	778	32.4	804	12	BG872805	BG872805 602793925
37	778	32.4	837	12	BG869312	BG869312 602788827
38	778	32.4	856	12	BG870147	BG870147 602793390
39	777	32.4	799	12	BG872144	BG872144 602792982
40	776	32.4	769	12	BG868463	BG868463 602785844
41	775	32.3	836	12	BG870266	BG870266 602791134
42	774	32.3	817	12	BG873398	BG873398 602794380
43	774	32.3	828	12	BG869811	BG869811 602789327
44	773	32.2	728	12	BG866113	BG866113 602784938
45	773	32.2	745	12	BG871565	BG871565 602790019

ALIGNMENTS

RESULT 1  
AK009012  
LOCUS

DEFINITION  
AK009012  
Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2300002L19:homolog to CHITOTRIOSIDASE PRECURSOR, full insert sequence.

ACCESSION  
AK009012.1 GI:12843547

VERSION  
HTC; CAP trapper.

KEYWORDS  
Mus musculus (strain:CS7BL/6J) adult male tongue cDNA to mRNA.

SOURCE  
clone:lib-RIKEN full-length enriched mouse cDNA library

ORGANISM  
Mus musculus

Mus musculus

Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1	Carninci, P. and Hayashizaki, Y.
REFERENCE	High-efficiency full-length cDNA cloning
AUTHORS	
TITLE	
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636

## REFERENCE

**AUTHORS** Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
**TITLE** Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
**JOURNAL** Genome Res. 10 (10), 1617-1630 (2000)  
**MEDLINE** 20499374

PUBMED

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akaiyama, J., Nishi, K., Kitzunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Matsuda, S., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Chatsura, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

## MEDLINE

11076861  
4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Nishinshi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Kishi, K., Kiyosawa, H., Kondo, S., Yamahata, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nitaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staab, F., Suzuki, R., Tomita, M., Wagner, D., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Maehina, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Db 473 AGCCTAAAGACTAAGAACCCCAAGCTCAAGACCTGTTAGCCGTTGGAGCGCTGGACCTTT 532  
Qy 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100  
Db 533 GGTATCCCAAGATTCACAGACATGTGTGGCCACCGCCAGCAACCGCAGACCTTTGTGAAG 592  
Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120  
Db 593 TCAGCCCTAAGTTTCTGGCGACTCAAGGTTTGTATGGCTTGACCTTGACCTGGAGTTC 652  
Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140  
Db 653 CCAGGTGGACGTGGGAGCCCAAGTAGACAAGAGAGATTCCACGCCCTGATACAGGAC 712  
Qy 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160  
Db 713 TTGGCCAAAGCCTTCAGGAGGAAGCCCAAGTCTCAGGGAAGGAACGCCCTCTCTGACT 772  
Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180  
Db 773 GCAGCTGTACCAAGTATCGAGCGCTGTGGATGTGGCTAGAGGTGGACAGATTGCC 832  
Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200  
Db 833 CAGAGCTTGATTTTCATCAACCTTATGGCTTACGACTTCACAGCTCTCTTGGAAGACC 892  
Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu 220  
Db 893 ACAGGGCATAAACAGCCCTCTACAAAGGCAAGGAGAAAGTGGGACACCCCTCAGCNA 952  
Qy 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240  
Db 953 AACGTGGATGCTGTGTGACGCTCTGGCTGCAAGAGGGACCCAGCCAGCAAACTGATC 1012  
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260  
Db 1013 CTGGCATGCTTACCTATGAGCGCTCTTTCACTTGGCTCTCTCGTCAGACAATGGAGTT 1072  
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280  
Db 1073 GGGGCCCCAGCACAGGGCTGTGGTCCCGAGGCCCTTATACGAAGCAAAAGGGGTCCTG 1132  
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGluArgIleGlnAspGlnLys 300  
Db 1133 GCTTACTATGAGCTGCTCTTGGAAG-----GAAAGACAGAAATCCAGGACCAAG 1186  
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320  
Db 1187 GTGCCTTACGCTTCCAGGACCAACCACTGGTGAGCTTTGACGACGTGGAAAGCTTCAAA 1246  
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLysGlyAlaMetValTrpAlaLeuAsp 340  
Db 1247 GC----- 1248  
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProGluLeuGlnThrLeu 360  
Db 1248 ----- 1248  
Qy 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuValProLys 380  
Db 1248 ----- 1248  
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400  
Db 1249 -----CAA 1251  
Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420  
Db 1252 GGCAAGCTGATGGGTCTACCCCAACCTCGAGACGAGTCCACTTACTACAACTGTGGA 1311  
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440

Db 1312 GGAGGCGCGCTGTTCCAGCAGAGCTGTCTCCAGCGCTGTGTGTTTAGAGCCCTCTTGCAAA 1371  
Qy 441 CysCysThrTrpAsn 445  
Db 1372 TGTGTACTCTGGAGC 1386  
RESULT 2  
LOCUS AK008633 1533 bp mRNA linear HTC 19-JAN-2002  
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:chitinase, acidic, full insert sequence.  
ACCESSION AK008633  
VERSION AK008633.1 GI:12842941  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library  
Clone:2200003E03.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4  
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bulc, C., Fletcher, C., Fujita, M., Gariboldi, M., Guetincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409 (6821), 685-690 (2001)  
MEDLINE 21085660  
PUBMED 11217851  
REFERENCE 5  
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,





```

putative"
/codon_start=1
/protein_id="BAB25878.1"
/db_xref="GI:12843149"
/db_xref="MG:1932052"
/translation="MQNNEITIEWNDVLYKAFNDLKNRNSKLTLAIGWNFFOTA
PFTTMYSTQNRQFTTTSVIFKRYQYGFQGLDLDWEYPGSRGSPQDQKHLFTVLVKEM
REAFEOEALIESNRPRLMTVAAGGINSIQAGVEIPELSKYLDPIHVMYVDLHGSWEG
VTGENSPLYKYPPTETGSNAYLVNDVYMYWKNNGAPAEKLIYVGFPEYGHAFILRNPDS
NLGATSGDGPAGPYTRQAGFWAYYEICTFLRSGATEVWDASQ"
polyA_signal
1520..1525
/notes="putative"
polyA_site
1535
/notes="putative"
BASE COUNT 379 a 381 c 373 g 402 t
ORIGIN

Alignment Scores:
Pred. No.: 8,07e-118 Length: 1535
Score: 1304.50 Matches: 236
Percent Similarity: 67.11% Conservative: 70
Best Local Similarity: 51.75% Mismatches: 129
Query Match: 54.40% Indels: 21
DB: 11 Gaps: 3

US-10-004-219B-10 (1-445) x AK008757 (1-1535)
QY 3 LeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22
DB 86 CTGATATGCTATTTCACCAACTGGGCCCGAGTATCGGCCAGGCTTGGGGAGCTTCAAGCCT 145
QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42
DB 146 GATGACATTAAACCCCTGCTGTGTACTACCTGATCTATGCCTTGTCTGGGATGAGAAC 205
QY 43 HisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
DB 206 AATGAGATCACCACCATAGATGAATGGAATGATGTTACTCTCTATAAAGCTTTCATGACTTG 265
QY 63 LysLeuMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPheGlyThr 82
DB 266 AAAAAAGGACGACGACCAATGAAACCCCTCGCAATGGAGCTGGGAGCTTGGAACTTGAAC 325
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
DB 326 GCTCCTTTTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACCTTCTATTACCTCAGTC 385
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 386 ATCAAAATTTCTGGCTCAGTATGATGGTTGTGATGGACTGGACTGGGAAATACCCAGGC 445
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
DB 446 TCACGTGGGAGGCCCTCCTCAGGACAAAGCATCTCTTCACTGTCCTGTGTGAAGAAATGCGT 505
QY 143 AsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAlaAla 162
DB 506 GAAGCTTTTGAGCAGAGGCTATTGAGAGCAACAGGCCAGAGCTGATGTTACTGCTGTCT 565
QY 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182
DB 566 GTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTGATGATGATCCCTGAACCTTCTTAAGTAC 625
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
DB 626 CTGGATTTTCATCATGTCATGATACATACCTCCATGCTGCTCTGGGAGGGCTACATGGG 685
QY 203 HisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyValAlaAlaSerLeuAsnVal 222
DB 686 GAGATAGTCTCTTTTACAAATACCTACTGAGACTGGTAGCAATGCCCTCACTCAATGTG 745
QY 223 AspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGly 242
DB 940

```

```

DB 746 GATTATGTCACTGAACCTATTGGGAAGAACCAATGGAGCCCGCCAGCTGAGAGCTATTGTGGA 805
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262
DB 806 TTCCAGAGTAGTGGACACGCCCTTTCATCTCGAAGAACCCCTCTGTAATGAATGTGTGCC 865
QY 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAlaTyr 282
DB 866 CTACTCTGTGTGATGGCCCTCTGGCCCTATACACAGACAGCTGGTCTTCTGGGCCTAC 925
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 926 TATGAGATTTCACCTTTCTGAGAAAGTGAGCACTGAGGTCCTGGGATGCCTCCCAATAA 985
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
DB 986 GTGCCCTATGCTTATAAGGCCAACGAGTGCTGTGCTATGACCAATATCAAGAGCTTCAGT 1045
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1046 GTTAAGGCTCAGTGGCTTAAGCAGAACCAATTTTGGAGGTGCCATGATCTGGGCCATTGAC 1105
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1106 CTTGATGACTTCACTGGCTCTTTCTGTGATCAGGAGAAATTTCTCTGACTTCTACTTTG 1165
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1166 AACAAAGCCCTTGGCATATPCCACTGAAGTTGCCAGCTCTCTGACGTG-----Glu 1213
QY 381 ProGlyGlnProSerGluPro-----CCTTCGAGCCAGTACTCTCTCAGGAAGTGGGAGTGGGAGC 1264
DB 1214 -----CCTTCGAGCCAGTACTCTCTCAGGAAGTGGGAGTGGGAGTGGGAGC 1264
QY 389 HisGlyProSerProGlyClnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro 408
DB 1265 TCCGGAGGAAGCTCTGGAGCGAGTGATTTCTGTGCCGACAAAGCAGATGGCCTCTACCT 1324
QY 409 AsnProArgGluArgSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428
DB 1325 GTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAAATCATAACAGCAGCAT 1384
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp 444
DB 1385 TGTCAAGCAGGCTGTTTGTGATACCACTGATGTAATTTGTCGCACTGG 1432

RESULT 4
AK008650 1534 bp mRNA linear HTC 19-JAN-2002
LOCUS Mus musculus adult male stomach cDNA, RIKEN full-length enriched
DEFINITION library, clone:2200008L17:chitinase, acidic, full insert sequence.
ACCESSION AK008650
VERSION AK008650.1 GI:12842971
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA,
clone lib:RIKEN full-length enriched mouse cDNA library
clone:2200008L17.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374

```





686 GA-AATAGTCCTCTTTTACAAATACCTACTGAGACTGGTAGCAATGCCTACCTCAATGTG 744  
QY 223 AspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242  
Db 745 GATTATGTCATCACTAATGTTGGAGAACAAATGGAGCCCGCTGAGAGCTCATGTGTGA 804  
QY 243 MetProThrTyxGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262  
Db 805 TCCGA-GAGTATGCACACACCTTCATCTGAGAAACCCCTCTGATATGAATGGTGGTGC 863  
QY 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAlaTyx 282  
Db 864 CTTACCTGCTGGTATGCGCTGCTGGCCCTATACACAGACAGCTGGTCTGGGCTAC 923  
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 924 TATGAGATTGACCTTCTTGAGAAAGTGGAGCCAGCTGAGGCTGGGATGCTCCCAAGAA 983  
QY 301 ValProTyxIlePheArgAspGlnTrpValGlyPheAspValGluSerPheLys 320  
Db 984 GTGCCCTATGCTATAAGCCACGAGTGGCTGGCTATGACATATCAAGAGCTTCAGT 1043  
QY 321 ThrLysValSerTyxLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340  
Db 1044 GTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCGGGCATTGAC 1103  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyxProLeuIleGlnThrLeu 360  
Db 1104 CTGTGATGACTTCACTGGCTTCTTCTGTGATCAGGAAATTTCTCTGACTTCTACTTTG 1163  
QY 361 ArgGlnGluLeuSerLeuProTyxLeuProSerGlyThrProGluLeuGluValProLys 380  
Db 1164 AACAAAGCCCTTGGCATATCCAGTGAAGTGGACAGCTCTGACGTG-----Glu 388  
QY 381 ProGlyGlnProSerGluPro-----CCTTCCGAGCAGTGAAGTGGAGTGGGGTGAAGC 1262  
Db 1212 -----CCTTCCGAGCAGTGAAGTGGAGTGGGGTGGAGTGGGGTGGAGC 1262  
QY 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyAlaAspGlyLeuTyxPro 408  
Db 1263 TCGGAGGAGCTCTGGAGCAGTGGATTTCTGCGCAGCAAGACAGATGCTCTACCT 1322  
QY 409 AsnProArgGluArgSerPheTyxSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428  
Db 1323 GTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCATACACAGCAGCAT 1382  
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp 444  
Db 1383 TGTCAAGCAGGGCTGTTTGTGATACCAAGCTGTAATGCTGCAACTGG 1430  
RESULT 5  
AK007573 LOCUS AK007573.1 GI:12841201  
DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length  
enriched library, clone:1810021D23;chitinase, acidic, full insert  
sequence.  
ACCESSION AK007573  
VERSION AK007573.1  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to  
Meth. Enzymol. 303, 19-44 (1999)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 9279253  
PUBMED 10349636  
AUTHORS 2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

QY 243 MetProThrTyxGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262  
Db 805 TCCGA-GAGTATGCACACACCTTCATCTGAGAAACCCCTCTGATATGAATGGTGGTGC 863  
QY 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAlaTyx 282  
Db 864 CTTACCTGCTGGTATGCGCTGCTGGCCCTATACACAGACAGCTGGTCTGGGCTAC 923  
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300  
Db 924 TATGAGATTGACCTTCTTGAGAAAGTGGAGCCAGCTGAGGCTGGGATGCTCCCAAGAA 983  
QY 301 ValProTyxIlePheArgAspGlnTrpValGlyPheAspValGluSerPheLys 320  
Db 984 GTGCCCTATGCTATAAGCCACGAGTGGCTGGCTATGACATATCAAGAGCTTCAGT 1043  
QY 321 ThrLysValSerTyxLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340  
Db 1044 GTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCGGGCATTGAC 1103  
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyxProLeuIleGlnThrLeu 360  
Db 1104 CTGTGATGACTTCACTGGCTTCTTCTGTGATCAGGAAATTTCTCTGACTTCTACTTTG 1163  
QY 361 ArgGlnGluLeuSerLeuProTyxLeuProSerGlyThrProGluLeuGluValProLys 380  
Db 1164 AACAAAGCCCTTGGCATATCCAGTGAAGTGGACAGCTCTGACGTG-----Glu 388  
QY 381 ProGlyGlnProSerGluPro-----CCTTCCGAGCAGTGAAGTGGAGTGGGGTGAAGC 1262  
Db 1212 -----CCTTCCGAGCAGTGAAGTGGAGTGGGGTGGAGTGGGGTGGAGC 1262  
QY 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyAlaAspGlyLeuTyxPro 408  
Db 1263 TCGGAGGAGCTCTGGAGCAGTGGATTTCTGCGCAGCAAGACAGATGCTCTACCT 1322  
QY 409 AsnProArgGluArgSerPheTyxSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428  
Db 1323 GTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCATACACAGCAGCAT 1382  
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp 444  
Db 1383 TGTCAAGCAGGGCTGTTTGTGATACCAAGCTGTAATGCTGCAACTGG 1430  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409 (6821), 685-690 (2001)  
MEDLINE 21085660  
PUBMED 11217851  
REFERENCE 5 (bases 1 to 1530)  
AUTHORS Arakawa, T., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,  
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanaoka, T.,  
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,  
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
Kurihata, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,  
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer





/clone\_lib="NCI CGAP Lu33"  
/tissue\_type="pooled lung tumors"  
/lab\_host="DH10B (phage-resistant)"  
/note="organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "  
BASE COUNT 202 a 248 c 230 g 157 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,766-84 Length: 837  
Score: 960.50 Matches: 185  
Percent Similarity: 79.38% Conservative: 19  
Best Local Similarity: 71.98% Mismatches: 48  
Query Match: 40.05% Indels: 5  
DB: 13 Gaps: 2

US-10-004-219B-10 (1-445) x B1412518 (1-837)

QY 191 TyrAspPheHisGlySerTrpGluLysValThr-GlyHisAsnSerProLeuTyrLysAr 210  
DB 5 TACGACTTCCAGCTCTCTGGAAAG---ACACAGGCATAACAGCCCTCTCAAAAG 61  
QY 210 ggInGluLysSerGlyAlaAlaSerLeuAsnValAspAlaAlaValGlnGlnTrpLe 230  
DB 62 GCAAGAGAAAGTGGGCGAGCGCTGAGCAAAACGTGGATGCTGTGACGCTCTGGCT 121  
QY 230 uGlnLysGlyThrProAlaSerLysLeuLeuLeuGlyMetProThrTyrGlyArgSerPh 250  
DB 122 GCAGAAAGGAGCCGCCAGCCAGCAAACTGATCTCTGGCATGCCCTACTATGGACGCTCTTT 181  
QY 250 eThrLeuAlaSerSerAspThrArgValGlyAlaProAlaThrGlySerGlyThrPr 270  
DB 182 CACCTTGGCTCTCTGTACAAATGAGTGGGCCCCCAGCCAGGCGCTGGTGGCCCC 241  
QY 270 oGlyProPheThrLysGluGlyMetLeuAlaTyrTyrGluValCysSerTrpLysGl 290  
DB 242 AGGCCCTATACGAAGGAGGAGGCTCTGCTTACTATGAGGCTGCTCTCGAAG-- 299  
QY 290 yAlaThrLysGlnArgIleGlnAspGlnLysValProTyrIlePheArgAspAsnGlnTr 310  
DB 300 ----GAAAGACACAGAAATCGAGGACAGAGGAGGCTTACGCGCTTCCAGGACCAACGAGTG 355  
QY 310 pValGlyPheAspAspValGluSerPheLysThrLysValSerTyrLeuLysGlnLysGl 330  
DB 356 GGTGAGCTTTCAGAGCTGGAAGGCTTCAAGCCAGGCTGCTTCTTCAACAGAGGG 415  
QY 330 yLeuGlyGlyAlaMetValTrpAlaLeuAspLeuAspPheAlaGlyPheSerCysAs 350  
DB 416 GCTGGGAGGAGCATGCTCTGGCTCTGGACTTGGATGACTTCAAGGCTTCTCTGCA 475  
QY 350 nGlnGlyArgTyrProLeuIleGlnThrLeuArgGlnGlnLeuSerLeuProTyrLeuPr 370  
DB 476 CCAGGCGCCGTACCTCTCATCCGACACTACGCGAGGAACCTAAATCTTCCATCCGAGAC 535  
QY 370 oSerGlyThrProGluLeuValProLysProGlyGlnProSerGluProGluHisGl 390  
DB 536 TCAAGGAGGCCAGAACAGATAATACCTGAGCCAGCCCATCTTCTATGCCAGAGCAGGG 595  
QY 390 yProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrProAsnPr 410  
DB 596 ACCAGCCAGGCTAGATAACTTCTGCCAAGGCAAGCTGATGGGTCTACCCCAACC 655  
QY 410 oArgGluArgSerSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSerCysPr 430  
DB 656 TGGAGACGAGTCCACTTACTACAACTGTGGAGGAGGGCGCTGTTCCAGCAGAGCTGTAC 715

QY 430 oThrGlyLeuVal-PheSerAsnSerCysLysCysCysThrTrpAsn 445  
DB 716 TCCAGGCTGTGGTTAGAGCCCTCGCAATGTGTACCTGGAGC 762

## RESULT 8

BQ231098

## LOCUS

BQ231098

## DEFINITION

AGNCOURT 7578138 NCI CGAP St1 Mus musculus cDNA clone

IMAGE:6051580 5', mRNA sequence.

BQ231098

BQ231098.1 GI:20412498

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13305 row: k column: 05

High quality sequence stop: 772.

Location/Qualifiers

1..892

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:6051580"

/clone\_lib="NCI CGAP St1"

/lab\_host="DH10B (T1-resistant)"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.77 kb. Library constructed by Life

Technologies. Note: this is a NCI CGAP Library."

I others

BASE COUNT 219 a 226 c 216 g 230 t

ORIGIN

Alignment Scores:

Pred. No.: 7,286-79 Length: 892

Score: 906.00 Matches: 161

Percent Similarity: 74.18% Conservative: 43

Best Local Similarity: 58.58% Mismatches: 71

Query Match: 37.78% Indels: 0

DB: 14 Gaps: 0

US-10-004-219B-10 (1-445) x BQ231098 (1-892)

QY 3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22

DB 42 CTGATATGCTATTTCCCAACTGGGCCAGTATCGCCAGGCTTCGAGGAGCTTCAAGCCT 101

QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42

DB 102 GATGACATTAACCCCTGCTGTACTCACCTGATCTATGCTTTGCTGGATGAGAAC 161

QY 43 HisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62

DB 162 AATGAGATCACCAACCATAGATGATGTTACTCTCTATAAAGCTTTCATGACTTG 221

QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnGlyThr 82

DB 222 AAAAAACAGAAACCAAACTGAAACCCCTCTGGCAATTGGAGGCTTGAACCT 281

QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102

```
Db 282 GCTCCTTTTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACCTTCATTACCTCAGTC 341
Qy 103 IleArgPheIeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 342 ATCAAAATTTTCGCGTCAGTATGGTGTGATGGAGCTGACCTGGAGTGGGAATACCCAGGC 401
Qy 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
Db 402 TCACGTGGAGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTGCTGGTGAAGNAATGCCT 461
Qy 143 AsnAlaPheGlnGlnGlnGlnGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162
Db 462 GAAGCTTTTTCAGCAGGAGGCTATTGAGAGCAACAGCCAGCCAGACTGATGTTACTGTCT 521
Qy 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysLysLeuAlaGlnAsn 182
Db 522 GTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACCTTTCTAAGTAC 581
Qy 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
Db 582 CTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 641
Qy 203 HisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222
Db 642 GAGAAATAGTCTCTTTTACAAATACCTCTAGAGACTGGTAGCAATGCCTACCTCAATGTG 701
Qy 223 AspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242
Db 702 GATTATGTCTAGTAACATTTTGGAGAAACAATGGAGCCCGAGCTGAGAGCTCATTTGTTGA 761
Qy 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262
Db 762 TTCCCAGATGATGGACACACTTCCTCTGAGAAACCCCTCTGATTAATGGAATGTTGCC 821
Qy 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGly 277
Db 822 CCTACCTCTGGTGTGATGGCCCTGCTGGCCCTATACCCAGACAGAGCC 866

RESULT 9
BG867815 838 bp mRNA linear EST 29-MAY-2001
LOCUS 602786336F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912544 5',
DEFINITION mRNA sequence.
ACCESSION BG867815
VERSION BG867815.1 GI:14218355
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 838)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M10816 row: 9 column: 09
High quality sequence stop: 824.
Location/Qualifiers
1. .838
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4912544"
/clone_lib="NCI_CGAP_SG2"
```

FEATURES  
source

```
BASE COUNT 209 a 211 c 198 g 220 t
ORIGIN
Alignment Scores: 2,876-74 Length: 838
Pred. No.: 859.00 Matches: 155
Score: 73.31% Conservative: 40
Best Local Similarity: 58.27% Mismatches: 71
Query Match: 35.82% Indels: 0
DB: 12 Gaps: 0

US-10-004-219B-10 (1-445) x BG867815 (1-838)
Qy 3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22
Db 40 CTGATATGCTATTTTCCCAACTGGGCCAGTATCGCCAGGTCCTGGGAGCTTCAAGCCT 99
Qy 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42
Db 100 GATGACATTAACCCCTGCTGTGTACTCACCTGATCTATGCCCTTTGCTGGGATGCAGAAC 159
Qy 43 HisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
Db 160 AATGAGATCACCACCATAGATGGAATGATGTTACTCTCTATAAAGCTTTCATGACTTG 219
Qy 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
Db 220 AAAACAGCAGCAGCAAACTGAAACCCCTCTGCAATTTGGAGGCTTGAACCTTTGGAAC 279
Qy 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
Db 280 GCTCCTTTTCACTACCATGCTTTTCCACTCTTCAGAACCCGACAGCTTCACTACCTCAGTC 339
Qy 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 340 ATCAATTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 399
Qy 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
Db 400 TCACGTGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTGCTGAGGAAATGCCT 459
Qy 143 AsnAlaPheGlnGlnGlnGlnGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162
Db 460 GAAGCTTTTTCAGCAGGAGGCTATTGAGAGCAACAGCCCGACAGCTGATGCTGCTGTCT 519
Qy 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysLysLeuAlaGlnAsn 182
Db 520 GTAGCTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACCTTCTAAGTAC 579
Qy 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
Db 580 CTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 639
Qy 203 HisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222
Db 640 GAGAAATAGTCTCTTTTACAAATACCTCTAGAGCTGGTAGCAATGCCTTACCTCAATGTG 699
Qy 223 AspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242
Db 700 GATTATGTCATGAACTATTGGAGAACATGGAGCCCGACAGCTGAGAGCTCATTTGTTGA 759
Qy 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262
Db 760 TTCCCAGAGTGTGGACACACACTTCTCTGAGAAACCCCTCTGATAATGGAATTTGTTGCC 819
Qy 263 ProAlaThrGlySerGly 268
Db 820 CCTACTCTGTTGTAATGGC 837
```

/lab host="DH10B (T1 phage-resistant)"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

```

RESULT 10
BQ231064
LOCUS      BQ231064
DEFINITION AGENCOURT 7578230 NCI CGAP_St1 Mus musculus cDNA clone
IMAGE:6051490 5', mRNA sequence.
ACCESSION  BQ231064
VERSION    BQ231064.1
KEYWORDS   BQ231064.1 GI:20412464
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 872)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13305 row: 9 column: 11
High quality sequence stop: 724.
Location/Qualifiers
1..872
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6051490"
/clone_lib="NCI CGAP_St1"
/lab_host="DH10B (TI-resistant)"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.77 kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT  218 a 222 c 205 g 225 t
ORIGIN
Alignment Scores:
Pred. No.:      1,41e-71      Length:      872
Score:          832.00      Matches:    154
Percent Similarity: 73.3%      Conservative: 39
Best Local Similarity: 58.5%      Mismatches: 68
Query Match:    34.70%      Indels:     2
DB:             14          Gaps:       0
US-10-004-219B-10 (1-445) x BQ231064 (1-872)
QY      3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      74 CTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGGAGCTTCAAGCT 133
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      23 LysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThrAsn 42
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      134 GATGACATTATACCCCTGCTGTGTACTACCTACCTGATCTATGCTCTTGGGATGAGAAC 193
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      43 HisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      194 AATGATACCAACCACTAGATGAATGATGATGATCTCTATTAAGCTTTCATGACTTG 253
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      254 AAAAAACAGAACACGACAACTGAAACCTCTCTGCGCAATTGGAGGTGGAACCTTGGAACT 313
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      314 GTCCTCTTTCACCTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCACTACCTCAGTGC 373
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122

```

```

      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      374 ATCAAAATTTCTGCGTCACTATGGGTTTCATGAGCTGGACCTGGAGTAATACCCAGGC 433
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      434 TCAGTGGGAGCCCTCTCCAGACAAGCATCTCTTCACTGTCTCTGGTGAAGAAATGGGT 493
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      143 AsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAlaAla 162
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      494 GAAGCTTTTGGAGGAGGCTATTGAGAGCAACAGGCCCCAGACTGATGGTTACTCTGCT 553
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      163 ValProAlaGlnGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      554 GTAGCTGTGGGATTTCCACATCCAGGCTGGCTATGATCCCTGAACTTTCTTAAGTAC 613
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      614 CTGGATTTTCATCCATGATGACATATGACCTCATGCTCTCTGGGANGGCTACTGCG 673
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      203 HisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      674 GAGATAGTCTCTTTACAATACCTACTGAGACTGGTAGCAATGCCCTACTCTCAATGTG 733
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      223 AspAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      734 GATTATGTCACTAACTATTGGAAGAACAAATGAGGCCCGAGCTGAGAAGCTCATTTCTTGA 793
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      243 MetProThrTyrGlyArgSer-PheThrLeuAlaSerSerSerSerAspThr-ArgValGlyA 262
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      794 TTCACAGATATGGACACACCTTCTCATCTGAGAAACCTCTTGATAATGGGAATGGTG 853
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      262 laPro 263
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D      854 CCCCC 858
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 11
BQ868741
LOCUS      BQ868741
DEFINITION 602787570F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4913800 5',
mRNA sequence.
ACCESSION  BQ868741
VERSION    BQ868741.1
KEYWORDS   BQ868741.1 GI:14219281
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 837)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10819 row: k column: 17
High quality sequence stop: 836.
Location/Qualifiers
1..837
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4913800"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dt. Average insert size 1.3 kb. Constructed by Life

```

Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 212 a 211 c 197 g 217 t  
ORIGIN

## Alignment Scores:

Pred. No.:	1.66e-71	Length:	837
Score:	831.00	Matches:	149
Percent Similarity:	74.51%	Conservative:	41
Best Local Similarity:	58.43%	Mismatches:	65
Query Match:	34.65%	Indels:	0
DB:	12	Gaps:	0

US-10-004-219B-10 (1-445) x BG868741 (1-837)

```
QY 3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 CTGATATGCTATTTCACCACTGGGCCAGTATCGCCAGGTCGGGAGCTTCAAGCCT 131
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 23 LysAspLeuAspProSerLeuCysThrHisLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 GATGACATTAAACCCCTGCTGTACTACCTGATCTATGCTTTGCTGGATGCAGAAC 191
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 43 HisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
192 AATGAGATCACCACCATAGATGAATGATGTTACTCTCTATAAAGCTTTCAATGACTTG 251
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 AAAAAAGGACAGCACTGAATAACCTCTGCGAATGGAGCTGGAACTTTGGAAT 311
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
312 GCTCCTTTCACTACCATGCTGTTTCCACTTCTCAGAACCGCCAGACCTTCAATACCTCAGTC 371
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
372 ATCAAAATTTCTGGCTCAGTATGGTTGATGGACTGGACTGGGAATACCCAGGC 431
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 TCACGTGGAGGCCCTCTCAGACAGCATCTCTTCACTGCTCTGCTGGAGGAATGCCT 491
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 143 AsnAlaPheGlnGlnGlnAlaGlnThrSerGlyGluArgLeuLeuLeuSerAlaAla 162
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
492 GAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGCCAGCATGATGTTACTGCTGCT 551
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 163 ValProAlaGlyGlnThrThrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
552 GTAGCTGGTGGGATTTCCACATCCAGGCTGGCTATGAGATCCCTGAACCTTCTAAGTAC 611
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
612 CTGGATTTTCATCATGTCATGATACATATGACCTCCTCATGCTCTGGAGGGCTACACTGG 671
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 203 HisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
672 GAGATAGTCTCTTTACAAATACCTACTAGACTGGTAGCAATGCCTACCTCAATGTG 731
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 223 AspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
732 GATTATGTCATGAACCTATTGGGAGAGCAACATGAGCCCGCCAGTGGAGCTCATTTGTTGA 791
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAsp 257
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
792 TTCCAGAGTATGGACACACTTCAATCCTGAGAAACCCCTCTGAT 836
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

## RESULT 12

BG872647

LOCUS

DEFINITION 602793705F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4924897 5',

mRNA sequence.

ACCESSION BG872647

VERSION BG872647.1 GI:14223187

KEYWORDS  
SOURCE  
ORGANISMEST.  
house mouse.  
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 763)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLAM10848 row: j column: 02  
High quality sequence stop: 760.  
Location/Qualifiers  
1..763  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4924897"  
/clone\_lib="NCI CGAP SG2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

## FEATURES

source

BASE COUNT 199 a 190 c 178 g 196 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.83e-70 Length: 763  
Score: 814.00 Matches: 150  
Percent Similarity: 74.80% Conservative: 37  
Best Local Similarity: 60.00% Mismatches: 63  
Query Match: 33.94% Indels: 1  
DB: 12 Gaps: 0  
US-10-004-219B-10 (1-445) x BG872647 (1-763)

```
QY 3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 CTGATATGCTATTTCACCACTGGGCCAGTATCGCCAGGTCGGGAGCTTCAAGCCT 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 23 LysAspLeuAspProSerLeuCysThrHisLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GATGACATTAAACCCCTGCTGTACTACCTGATCTATGCTTTGCTGGATGCAGAAC 120
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 43 HisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 AATGAGATCACCACCATAGATGGAATGATGTTACTCTCTATAAAGCTTTCAATGACTTG 180
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 AAAAAAGGACAGCACTGAATAACCTCTGCGAATGGAGCTGGAACTTTGGAAT 240
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 GCTCCTTTCACTACCATGCTGTTTCCACTTCTCAGAACCGCCAGACCTTCAATACCTCAGTC 300
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 ATCAAAATTTCTGGCTCAGTATGGTTTGGAGCTGGACTGGGAATACCCAGGC 360
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 TCACGTGGAGGCCCTCTCAGACAGCATCTCTTCACTGCTCTGCTGGAGGAATGCCT 420
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```





Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10840 row: 1 column: 19  
High quality sequence stop: 845.

## FEATURES

source  
1. .880  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4921890"  
/clone\_lib="NCI CGAP S2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:  
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 218 a 220 c 211 g 231 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,34e-69 Length: 880  
Score: 808.00 Matches: 151  
Percent Similarity: 71.22% Conservative: 42  
Best Local Similarity: 55.72% Mismatches: 78  
Query Match: 33.69% Indels: 1  
DB: 12 Gaps: 0

US-10-004-219B-10 (1-445) x BG871468 (1-880)

QY 3 LeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyClnAlaArgPheLeuPro 22  
Db 67 CTGATATGTTATTCACCACTGGGCCAGTATCGGCCAGCTCTGGGAGCTTCAAGCCT 126  
QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42  
Db 127 GATGACATTAACCCCTGCTGTGTATCTACCTGATCTATGCTTCTGCTGGGATGAGAAC 186  
QY 43 HisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62  
Db 187 AATGAGATCACCACCATAGATGGAATGATGTTACTCTCTATAAGCTTCAATGACTTG 246  
QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyClyTrpAsnGlyThr 82  
Db 247 AAAAAAGGAGAACAGCAAACTGAAACCCCTCTGGCAATTTGGAGGCTGGAACTTTGGA 306  
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102  
Db 307 GCTCCTTTTCACTACCATGGTTTCCATCTTCAGAACCGCCAGACCTTCATCTACCTCAGTC 366  
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122  
Db 367 ATCAAAATTTCTGCTCAGTATGGTTTGTATGAGTCTGACCTGAGCTGGGAATACCCAGGC 426  
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrLeuValGlnAspLeuAla 142  
Db 427 TCACGTGGGAGCCCTCTCAGGACAAAGCATCTCTCAGTCTGCTGGTGAAGGAATGCGT 486  
QY 143 AsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162  
Db 487 GAAGCTTTTGAGCAGAGGAGGCTATTGAGCAACAGCCAGCTGATGGTTACTGCTGCT 546  
QY 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182  
Db 547 GTAGCTGGTGGGATTTCCAACTCCAGGCTGGCTGATGAGATCCCTGCACTTCTTAAGTAC 606  
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpClnLysValThrGly 202  
Db 607 CTGGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTCTGGGAGGCTACACTGGG 666

QY 203 HisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsnVal 222  
Db 667 GAGATAGTCTCTTTTACAAATACCTACTGAGACTGGTAGCAATGCCTACCTCAATGTG 726  
QY 223 AspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242  
Db 727 GATTATGTGATGACTATTGGGAAGAACCAATGGAGCCCAAGAGGTTCATTGCTGGAT 786  
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGlyAla 262  
Db 787 TCCAAGAGTATGGTGAACACCTTCATCTGAGAACCTCTGATACTGG-ATTGCTGCC 845  
QY 263 ProAlaThrGlySerGlyThrProGlyProPhe 273  
Db 846 CCTAATCTGCTGGTGGGCTGTGGGCTATTTC 878

## RESULT 15

BM458470

LOCUS

DEFINITION

BM458470

VERSION

BM458470.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12126 row: k column: 21

High quality sequence stop: 683.

Location/Qualifiers

1. .1098

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5497100"

/clone\_lib="NIH MGC 85"

/tissue\_type="lymphoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.867 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH MGC Library."

BASE COUNT 282 a 274 c 276 g 264 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 7,44e-69 Length: 1098

Score: 806.00 Matches: 162

Percent Similarity: 71.43% Conservative: 58

Best Local Similarity: 52.60% Mismatches: 75

Query Match: 33.61% Indels: 16

DB: 13 Gaps: 5

US-10-004-219B-10 (1-445) x BM458470 (1-1098)

QY 66 AsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrGlnLysPhe 85

Db 4 AATCCCAACTGAAATTTCTTTGTC-ATTGAGGGAGCTG-TTTGGTTCCAAAGGTTTC 61

```
QY 86 ThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAlaIleArgPhe 105
Db ||||| : : : : : |||||
62 CACCCTATGTTGATCTCTACATCACGCTTGGATTCATTAACTCCATAATCCTGTTT 121
QY 106 LeuArgLysTyrSerPheAspGlyLeuAspLeuAspTyrGluTyrProGlySerGlnGly 125
Db ||||| : : : : : ||||| : : : : : |||||
122 CTGAGGAACCATAACTTTCATGAGCTGGATTAAGCTGGATCTAGCCAGATCAGAAA--- 178
QY 126 SerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAlaAsnAlaPhe 145
Db ||||| : : : : : ||||| : : : : : |||||
179 -----GAAACACTCATTTCACTGCTGCTGATTCATGAGTTAGCAGAAGCCTTT 226
QY 146 GlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAlaValProAla 165
Db ||||| : : : : : ||||| : : : : : |||||
227 CAGAAGGACTTCACAAATCCACCAAGGAAGGCTCTCTGACTCGGGCGTATCTGCA 286
QY 166 GlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsnLeuAspPhe 185
Db ||||| : : : : : ||||| : : : : : |||||
287 GGGAGGCAAAATGATTGATAACAGCTATCAAGTTGAGAAACTGCAAAAGATCTGGATTTC 346
QY 186 ValAsnLeuMetAlaTyrAspPheHisGlySerTyrGluLys-----ValThrGlyHis 203
Db ||||| : : : : : ||||| : : : : : |||||
347 ATCAACCTCCTGCTCTTCACTTCCATGGGTCTTGGGAAAGCCCTTATCACTGCCCAC 406
QY 204 AsnSerProLeuTyrLysArgGlnGlnGluSerGlyValAlaAlaSerLeuAsnValAsp 223
Db ||||| : : : : : ||||| : : : : : |||||
407 AACAGCCCTCTGAGCAAGGGGTGGCAGGACAGAGGGCCCAAGCTCTACTACATGTGGAA 466
QY 224 AlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGlyMet 243
Db ||||| : : : : : ||||| : : : : : |||||
467 TATGCTGTGGGTACTGGGATACATAAGGGAATGCCATCAGAGAAGGTGTCATGGGCATC 526
QY 244 ProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAlaPro 263
Db ||||| : : : : : ||||| : : : : : |||||
527 CCCACATATGGGCACCTCTCACATGGCC---TCTGCAGAAACCAACCGTGGGGGCCCT 583
QY 264 AlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAlaTyrTyr 283
Db ||||| : : : : : ||||| : : : : : |||||
584 GCCTCTGGCCCTGGAGCTGCTGACCCCATCACAGAGTCTTCAGGCTTCCTGGCCTATTAT 643
QY 284 GluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAspGlnLysValPro 302
Db ||||| : : : : : ||||| : : : : : |||||
644 GAGATCTGCCAGTTCCTGAAAGAGCAAGATCACGCGGCTCCAGGATCAGCAGGTTCC 703
QY 303 TyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThrLys 322
Db ||||| : : : : : ||||| : : : : : |||||
704 TACGCAGTCAAGGGGAACCAAGTGGGTGGCTATGATGTGAAGAGTATGGAGAACAG 763
QY 323 ValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeuAsp 342
Db ||||| : : : : : ||||| : : : : : |||||
764 GTTCAGTCTTAAAGAT-TTAAACCTGGGAGGAGCCATGATCTGGTCTATTGACATGGAT 822
QY 343 AspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArgGln 362
Db ||||| : : : : : ||||| : : : : : |||||
823 GACTTCACCTGGCAAAATCCTGCAACCAAGGCCCTTACCTCTTTGTTCAGCAGTCAAGAAAG 882
QY 363 GluLeuSerLeuProTyrLeuPro 370
Db ||||| : : : : : |||||
883 -----CCTTGGCTCCCT 894
```

Search completed: June 30, 2003, 01:07:48  
Job time : 1478.66 secs

**THIS PAGE BLANK (USPTO)**



[illegible]







Db 499 FTCPPGTFLDPAALHCNW 516

## RESULT 10

A33918  
chitinase (EC 3.2.1.14) precursor - braconid wasp (*Chelonus* sp.)  
C:Species: *Chelonus* sp.  
C>Date: 28-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 21-Jul-2000  
R:Krishnan, A.; Nair, P.N.; Jones, D.  
J. Biol. Chem. 269, 20971-20976, 1994  
A:Title: Isolation, cloning, and characterization of new chitinase stored in active form  
A:Reference number: A53918; MUID:94342256; PMID:8063715  
A:Accession: A53918  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-483 <KRA>  
A:Cross-references: GB:U10422; NID:g533504; PIDN:AAA61639.1; PID:g533505  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 36.1%; Score 864.5; DB 2; Length 483;  
Best Local Similarity 37.6%; Pred. No. 1.9e-57;  
Matches 178; Conservative 79; Mismatches 171; Indels 45; Gaps 13;  
QY 2 KLVCYFTNWAQYRGGEARFLPKDLPSCLTTHLIYAFAGMTNHQSLSTTE-WND-ETLY 56  
Db 23 KVCYFGAWSVYRQGNKFDINGIDPTLCTHLYISFVGNGDKVVLDPWSLPGNLDGF 82  
QY 57 QEFNGLKKNPKLTLLAIGGNWFGTKQFTDMVATANNROTTFVNSAIRFLRYKSFDFGLD 116  
Db 83 GKFTSLRKNPKSVKIMAVGNGWAGSVFSPQASDQATREAFQANVVKFLOQYQDFDI 142  
QY 117 DWEYPGSQSPAVDKERFTTLVQDLANAFQOEAQTSKGERLLLSAAVPAGOTVYVDAGYEV 176  
Db 143 DWEYPAQRGSGPADVKNWKLCKALKAPVQ-----HDYLSRAVAAPETASAKSDI 195  
QY 177 DKIAQNLDVNLMAFYDFHGSWEKVTGHSNPLYKROBESGAALSNVDAVQOVLQKGTTPA 236  
Db 196 AEMSVYLDIFNLMTYDFHGPWBGHTGMHAPSSASHDSGNELKLVNKAAYKWLQNGVPK 255  
QY 237 SKLIIGMPTYGRSFTLASSSDTRVGAPATGSGTPGPFYKGGMLAYYEVCSW-KGATQOR 295  
Db 256 EKLVGVGPAYGKSFSLNSPNKGLGAPVSGAGTAGPYTGENGLLGYNETCEMKGAGDMEV 315  
QY 296 IDOKV-VPIFRDNQWVGDFDVEFKTKVSLKQKGLGGAMVWALDLDLDFAGFSCNOGR 353  
Db 316 VDNEKGVPIYAVKQWVSFDDLAIAKAQAKIYKQEGLGAMVWETDFFKGL-CGE-K 373  
QY 354 YPLIQTLRQEL-----SLP-----YLPSTGTP-----LEVPKQGPSEPHGSPGQD 396  
Db 374 YPVLKALNSVLGRGSSSPAETKRKNVPPDQAPPRSPAESDAPAPVPEVPSSESGE- 432  
QY 397 TFCQKADGLYNPRERSFYSC-----AAGRLFOQSCPTGLVFNNSCKCTW 444  
Db 433 --CSSVGQPLV---GQNCGLVCDGDMGGRKIPQVCPQGLCFNPANNYCDW 480

## RESULT 11

A38221  
chitinase (EC 3.2.1.14) Mf1 - nematode (*Brugia malayi*)  
C:Species: *Brugia malayi*  
C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
R:Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Plessens, W.F.; Perlier, F.B.  
Proc. Natl. Acad. Sci. U.S.A. 89, 1548-1552, 1992  
A:Title: Transmission-blocking antibodies recognize microfilarial chitinase in brugian  
A:Reference number: A38221; MUID:92179220; PMID:1542646  
A:Accession: A38221  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid; protein  
A:Residues: 1-504 <RUH>  
A:Cross-references: GB:M73689; NID:g156063; PIDN:AAA27854.1; PID:g156064  
A:Note: sequence extracted from NCBI backbone (NCBIP:85345)

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 35.9%; Score 860.5; DB 2; Length 504;  
Best Local Similarity 37.0%; Pred. No. 4e-57;  
Matches 177; Conservative 80; Mismatches 170; Indels 51; Gaps 9;  
QY 5 CYFTNWAQYRGGEARFLPKDLPSCLTTHLIYAFAGMTNHQSLSTT---EWNDET-----L 55  
Db 27 CYFTNWAQYRGGEKFLPGNIPNGLCTHLYIAFAKV--DELGDSKPFEMWEDTWSKGM 84  
QY 56 YQEFNGLKKNPKLTLLAIGGNWFGTKQFTDMVATANNROTTFVNSAIRFLRYKSFDFGLD 115  
Db 85 YSAVTKLRETNPLGLVLLSYGYNFGSAIFGIAKSAQKTERFKSAIAFLRKNFDFGD 144  
QY 116 LDWEYPGSQSPAVDKERFTTLVQDLANAFQOEAQTSKGERLLLSAAVPAGOTVYVDAGYEV 175  
Db 145 LDWEYFVGVV-----BEHAKLVEMKTAFAVEAKTSKQRLLLTAAVSAGKGTIDGSYN 198  
QY 176 VDKIAQNLDVNLMAFYDFHGSWEKVTGHSNPLYKROBESGAALSNVDAVQOVLQKGTTP 235  
Db 199 VESLGKPNFDLLFLMSYDLHGSWEKNVDLHGKLPHTKGEVSGIGIFNTEFAADYWASKGMP 258  
QY 236 ASKLIIGMPTYGRSFTLASSSDTRVGAPATGSGTPGPFYKGGMLAYYEVCSW-KGATQ 294  
Db 259 KEKIIIGIPMYAQQWTLDPNSETAIGAAASRPSSASKTNPAGGTASYWEICKYLKEGK 318  
QY 295 RIQOKV-PYIFRDNQWVGDFDVEFKTKVSLKQKGLGGAMVWALDLDLDFAGFSCNOGR 353  
Db 319 TVHGEVGAIVKGGQWGYDNEETIRIKMWLKEKGGYGFIAWALDFFDTGKSCGKGP 378  
QY 354 YPLIQTLRQELSLPYLPSTGTPLEVPKPG-----TEATEYDETEETSEAYDTDETEETSE 386  
Db 379 YPLLNALISSELE---GESENFETTEPSITETEAETDETEETSEAYDTDETEETSE 435  
QY 387 PEHGPSQDFTFCQ--GKADGLYNPRERSFYSCAAGRLFOQSCPTGLVFNNSCKCC 442  
Db 436 TEATTYDTDETEGECPEPDRGLFPHTDCHLFIQCANNIAYVMQCPATFFNDIAIKVC 493

## RESULT 12

A56596  
chitinase (EC 3.2.1.14) - tobacco hornworm  
C:Species: *Manduca sexta* (tobacco hornworm)  
C>Date: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #text\_change 29-Jan-1999  
R:Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.  
Insect Biochem. Mol. Biol. 23, 691-701, 1993  
A:Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinase  
A:Reference number: A56596; MUID:93357793; PMID:8353525  
A:Accession: A56596  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-554 <KRA>  
A:Cross-references: GB:U02270; GB:S64757; NID:g406048; PID:g406049  
A:Experimental source: larvae  
A:Note: sequence extracted from NCBI backbone (NCBIN:136417, NCBIP:136418)  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 34.4%; Score 824; DB 2; Length 554;  
Best Local Similarity 33.7%; Pred. No. 2.6e-54;  
Matches 181; Conservative 87; Mismatches 159; Indels 110; Gaps 15;  
QY 1 AKLVCFYTNWAQYRGGEARFLPKDLPSCLTTHLIYAFAGMTNHQSLSTTEWMDTEL----- 55  
Db 23 ARIVCFYNWAVYRQGVRYGIEDIPVEKCTHIIYSFIGV-----TEGSEVLIIDPE 75  
QY 56 -----YQEFNGLKKNPKLTLLAIGGNWFGTKQFTDMVATANNROTTFVNSAIRFLRK 108  
Db 76 LDVDKNGFRNTSLRSHSPSVKFWAVGVGWAEGSSKYSYHVAQKSTRMSFIRSVVSFLKK 135  
QY 109 YSPFDGLDWEYPGS--OGSPAVDKERFTTLVQDLANAFQOEAQTSKGERLLLSAAVPAG 166  
Db 136 YDFDGLDWEYPGGAADRGGSFSDKDFLYLVQELRRAFIRVGK--GWE---LTAAPFLA 190

```
QY 167 QTVDAQYEVKIAQNLDFVNLMAVDFHGSWEKVTGHNPSLYKROBESGAASLNVDAAV 226
Db 191 NFRLEMEGYHVPCLCOELDAIHVMSYDLRGNWAGFADVHSPLYKRPDHWAYEKLNVNDEL 250
QY 227 QOWLQGTSPASKLILGMPYGRSFTLASSD-----TRVGAPATGSGTGPFTKGGMLA 281
Db 251 HLWEEKGCPNKLWVGIPIFYGRSFTLSAGNNYVGLGTFINKEA-GGGDPAPYTNATGFWA 309
QY 282 YVEVCSW-----KGATKQRIQODQKVPYIFRDQWVGFDDVESFKTVSVLKQKGLGGAMV 336
Db 310 YVEICTEVKDQSGWTKWDEQCKPYAKYGTQWVGVEDQPSVEIKQWIKQKGYLGAMT 369
QY 337 WALDLDDFAGFSCNQGRYPLIOTLRELSPYLP-----SGTPELVKPKQCPSEPSH 389
Db 370 WALDDDFDQGL-CGE-KNPLIKILKHMSYVPPHTENTYPTPEWARP-FSTPSDPSE 426
QY 390 G-----PSPGQDT----- 397
Db 427 GDIPIPTTTAKPASTTKTTVKTITTTTAKPQSVIDEENDINVRPEKPEPEVEVP 486
QY 398 -----FCQKADGLYPNPRSSYFSCNAGRLFOQSCPTGLVFSNCKCTW 444
Db 487 PTNEVDGSEICNSDQDYI-PDKKHCDKYRWCNGEAMQFSCQHGTVFNVNLCVDM 542

RESULT 13
T14075
Chitinase (EC 3.2.1.14) - yellow fever mosquito
C:Species: Aedes aegypti (yellow fever mosquito)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14075
R:De la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A:Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.
A:Reference number: Z17872
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Accession: T14075
A:Molecule type: DNA
A:Residues: 1-1635 <DEL>
A:Cross-references: EMBL:AF026492; NID:g2564720; PID:g2564721; PIDN:AAB81850.1
C:Genetics:
A:Gene: CHT2
A:Introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 32.8%; Score 787.5; DB 2; Length 1635;
Best Local Similarity 41.4%; Pred. NO. 7e-51;
Matches 157; Conservative 72; Mismatches 121; Indels 29; Gaps 11;

QY 2 KLVCFYTNWAQYRQGEARFLPKDLPSLCTHLYAFAGTNGHQLST---TEWN--DETLY 56
Db 1258 KVCYFYNWATRWQGGKVTPTDDISNLCTHIVYGFAVLDRSLTIKTHDSWADINQFY 1317
QY 57 QFNGLKGNPKLKTLLAIGWNFGT-QKFTDMVATANNROTTFVNSAIRPKYISPDGLD 115
Db 1318 ERVVEHKKRGTKV--TLALGWNDSLGDKYSLVRSSARRAFVHAEFIEKYNFDGLD 1375
QY 116 LDWEYP-----GSQSPAVDKERFTTHVQDLANAFQGEAQTSKGERLLISAAPAGQTY 169
Db 1376 LDWEYFVQVQVCKGFPD-EKEGFAELVKELAEFR-----PRKLLSAAVSPSKV 1427
QY 170 VDAGYEVKIAQNLDFVNLMAVDFHGSWEKVTGHNPSLYKROBESGAASLNVDAAVQW 229
Db 1428 IDAGYDVPVLAIEYFDVIAWMTYDFHGHNDKQTHGVAFLYYPGDT--YDYFNANFSINTW 1485
QY 230 LQGTSPASKLILGMPYGRSFTLASSDTRVGAPATGSGTGPFTKGGMLAYEVC--- 286
Db 1486 IEKGAPSKRLVMGMPYGGQSFSLADSGKNGLNDKSYGPGEAGQFTRAGGLAFAYEICEKV 1545
QY 287 SWGATKQRIQODQKVPYIFRDQWVGFDDVESFKTVSVLKQKGLGGAMWALDDDFEA 345
Db 1546 NQNGNAVYRDSGRIGFYALRGQWVSVDVDEIRRSQFVKKNLGGGMWALDDDFR 1605
```

```
QY 346 GFSCNQGRYPLIOTLREL 364
Db 1606 G-RCGCGRHPLLKTMNQEL 1623

RESULT 14
A38368
Chitinase (EC 3.2.1.14) precursor - Bacillus circulans
C:Species: Bacillus circulans
C>Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 15-Oct-1999
C:Accession: A38368
R:Watanabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.
J. Biol. Chem. 265, 15659-15665, 1990
A:Title: Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evolutic
A:Reference number: A38368; MUID:90368776; PMID:2203782
A:Accession: A38368
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-699 <WAT>
A:Cross-references: GB:M57601; GB:J05599; NID:g1066341; PIDN:AAA81528.1; PID:g142688
C:Superfamily: fibronectin type III repeat homology
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 24.4%; Score 584.5; DB 2; Length 699;
Best Local Similarity 33.0%; Pred. NO. 4.5e-36;
Matches 147; Conservative 71; Mismatches 143; Indels 85; Gaps 16;

QY 2 KLVCFYTNWAQYRQGEARFLPKDLPSLCTHLYAFAGM----- 40
Db 45 KIVGYPSMAAYGR---NYNVADIDPTKVTHTINAFADICWNGIHGNDPDPGPNVTWTC 101
QY 41 TNHQLSTTE-----WND-----ETLYQ-----EFNGLKMKPKLKTLLAI 75
Db 102 QNEKQNTINVPNGTIVLGDPMIDTGTAGTWDQPIAGNINQLNKLQTNENLKTIIISV 161
QY 76 GGNFGTKQFTDMVATANNROTTFVNSAIRPKYISFDGLDWEYPSGQ-----SPAVD 130
Db 162 GGTW-SNRFSDVAATAATREVFANSVDFKRVNFDGVDLDWEYFVSGGLDGNKRPE 220
QY 131 KERFTTLVQDLANAFQGEAQTSKGERLLISAAPAGQTYVDAGYEVKIAQNLDFVNLMA 190
Db 221 KONYTLILSKIREKLDAAAGVDGK-KYLLTIASGASATYA-ANTELAIAIVDWINMT 278
QY 191 YDFHGSWEKVTGHNPSLYKROBESCA-----AASLNVDAAVQWLOKQTPASKLILGMPY 246
Db 279 YDFNGAWQKISAHNAPLVNDPAASAAGVDPDANTFNVAAGAQHLDAGVPAKLVGVPPY 338
QY 247 GRSFTLASSDTRVGAPATGSGTGPFTKGGMLAYEV---CSWKGATKQRIQODQKVP 302
Db 339 GRGWDGCAQAGNGQYQCTCTGGSSVG--TWEAGSFDYDLEANYINKNGYTRYWNTAKVP 396
QY 303 YIFR--DNQWVGFDDVESFKTVSVLKQKGLGGAMWALDDDDFAGFSCNQGRYPLIOTL 360
Db 397 YLYNASNKRFTSYDDAESVGYKTAIKSKLGGAMFWELSGD-----RNKTL 443
QY 361 ROELSLPLPSG--TPELEVPKPGQP 384
Db 444 QNKLKAD-LPTGTVPPVDTTAPSV 468

RESULT 15
D83764
Chitinase BH0916 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D83764
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D83764
A:Status: preliminary
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:51:08 ; Search time 7.89941 Seconds  
(without alignments)  
2336.500 Million cell updates/sec

Title: US-10-004-219b-10  
Perfect score: 2398  
Sequence: 1 AKLVCFYFTNWAQYRQGEARF.....QQSCPTGLVFSNCKCKCTWN 445

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwisProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1045.5	43.6	390	1 C312 HUMAN	Q15782 homo sapien
2	1041.5	43.4	383	1 C311 HUMAN	P36222 mus sapien
3	984.5	41.1	381	1 C311 MOUSE	Q61362 mus musculus
4	964.5	40.2	537	1 OGP BOVIN	Q28042 bos taurus
5	959.5	40.0	539	1 OGP SHEEP	Q28542 ovies aries
6	956	39.9	617	1 CHIT CAEEL	Q11174 caenorhabdi
7	947.5	39.5	527	1 OGP PIG	Q28990 sus scrofa
8	942	39.3	721	1 OGP MOUSE	Q62010 mus musculus
9	927	38.7	671	1 OGP MESAU	Q60557 mesocricetu
10	925.5	38.6	623	1 OGP PAPAN	P36718 papio anubi
11	917.5	38.3	678	1 OGP HUMAN	Q12889 homo sapien
12	860.5	35.9	504	1 CHIT BRUMA	P29030 brugia mala
13	824	34.4	554	1 CHIT MANSE	P36362 manduca sex
14	584.5	24.4	699	1 CHIT BACCI	P20533 bacillus ci
15	493.5	20.6	423	1 CH11 APHAL	P32470 aphanoclad
16	478	19.9	423	1 CH14 TRIHA	P48827 trichoderma
17	425.5	17.7	427	1 CH11 COCIM	P54196 coccidioid
18	422	17.6	499	1 CH1B SERMA	P11797 serratia ma
19	357	14.9	550	1 CHIT NPVOP	Q10363 orgyia pseu
20	355.5	14.8	563	1 CH1A SERMA	P07254 serratia ma
21	331.5	13.8	619	1 CHIT STRLI	P36909 streptomyce
22	331.5	13.8	820	1 CH1A ALTOS	P32823 alteromonas
23	329	13.7	551	1 CHIT NPVAC	P41684 autographa
24	328	13.7	1046	1 CH1D VIBFU	P96156 vibrio furn
25	305	12.7	610	1 CHIT STRPL	P11220 streptomyce
26	270.5	11.3	1146	1 KTXA KLJLA	P09805 kluveromyce
27	161	6.7	367	1 DIAC RAT	Q101460 rattus norv
28	155	6.5	427	1 YAAH BACSU	P37531 bacillus su
29	147.5	6.2	385	1 DIAC HUMAN	Q01459 homo sapien
30	127.5	5.3	491	1 YKCS CAEEL	Q14596 caenorhabdi
31	122	5.1	2334	1 WAPA BACSU	Q07833 bacillus su
32	120.5	5.0	524	1 CH1D BACCI	P27050 bacillus ci
33	113	4.7	1026	1 BGAL STRTR	P23989 streptococ

RESULT 1  
C312 HUMAN  
ID C312 HUMAN STANDARD; PRT: 390 AA.  
AC Q15782; Q15783; Q15749;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein 39).  
GN CH13L2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).  
RA Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R., Mak.T.W.;  
RT "Cloning of a novel lymphoid restricted human chitinase and localization to lp13.3".  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
RC TISSUE=Articular cartilage;  
RX MEDLINE=96325055; PubMed=8702629;  
RA Hu B., Trinh K., Figueira W.F., Price P.A.;  
RT "Isolation and sequence of a novel human chondrocyte protein related to mammalian members of the chitinase protein family".  
RL J. Biol. Chem. 271:19415-19420(1996).  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHOWN HERE) AND A LONG FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN, PANCREAS, AND LIVER.  
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcement/ or send an email to license@sib-sib.ch).

EMBL; U58515; AAB04534.1; -;  
EMBL; U58514; AAB04533.1; -;  
EMBL; U49835; AAC50597.1; ALT\_INIT.  
MIM; 601526; -;  
InterPro; IPR001579; Chitinase 18/2.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR PROSITE; PS01095; CHITINASE\_18; FALSE NEG.  
KW Hydrolase; Signal; Alternative splicing.  
FT SIGNAL 1 26  
POTENTIAL.

34 111.5 4.6 597 1 CHIX STROI  
35 108 4.5 36 1 C311\_BOVIN  
36 101.5 4.2 958 1 AMVG\_DEBOC  
37 101.5 4.2 1289 1 CSAB\_BACUD  
38 101.5 4.2 2336 1 POLG\_FMDVZ  
39 101 4.2 4639 1 DYHC\_DROME  
40 100 4.2 1313 1 ACE\_RAT  
41 99 4.1 293 1 CHIA\_CICAR  
42 99 4.1 502 1 WASP\_HUMAN  
43 99 4.1 1902 1 PIP\_LACLC  
44 98 4.1 956 1 NUTI\_MAGGR  
45 97.5 4.1 346 1 APBE\_HAEIN

ALIGNMENTS

Q05638 streptomyce  
P30922 bos taurus  
P22861 debaryomyce  
Q45753 bacillus th  
P43303 f genome po  
P37276 drosophila  
P47820 rattus norv  
P36908 cicer ariet  
P42768 homo sapien  
P16271 lactococcus  
Q01168 magnaporthe  
P44550 haemophilus

```
FT CHAIN 27 390 CHITINASE 3-LIKE PROTEIN 2.
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 1 12 MGATWDQKSLW -> SQQLWLGSPDEVWLYLVCETHFT
SQ SEQUENCE 390 AA; 43500 MW; 97B86A2F3AA35677 CRC64;
Query Match 43.6%; Score 1045.5; DB 1; Length 390;
Best Local Similarity 52.2%; Pred. No. 3.8e-72;
Matches 191; Conservative 76; Mismatches 90; Indels 9; Gaps 4;
QY 2 KLVCFYTNWAQYRQGEARFLPKDLPDLCTHLIYAFAGMTHNQLSTTEWDETLYQEFNG 61
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 KLVCFYTNWSQDRQEPCKFTPENIDPFLCSHLIYSPASINENKVIIDKSEVMVLYQTINS 87
QY 62 LKQWPKLTLIAIGWNGFTQKFTDMVATANNROTIVNSAIRFLKYSFDCGLDLDWEYP 121
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 LKTNPKLTLIAIGWNGFTQKFTDMVATANNROTIVNSAIRFLKYSFDCGLDLDWEYP 147
QY 122 GSQSPAVDKERFTTLVQDLANAFQQAOTSGKERILLLSAAVPAQOTYVDAGYEVDKIAQ 181
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 DQK-----ENTHETVLIELHAEAFQKDYKSTKERILLTAGVSAGRQMDINSYQVEKLAK 202
QY 182 NLDVNLMAYPHGSWEK--VTGNSPLYKROBESGAAASLNVDAAVQWLOKGTSPASKL 239
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 DLDVNLMAYPHGSWEKPLITGHNSPLSKGWDGPGSYNVVEYAVGIWIKGMPSEKV 262
QY 240 ILGMPYGRSFTLASSSDTRVGPATGSGTGPFTKEGGMLAYEVCWS-KGATKORIOD 298
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 VMGIPTVGHSTFLA-SAEFTVGPAPGPGGAAGPITESSGFLAYEICQFLKGAKITRLOD 321
QY 299 QKVPYIFRDNQWGFDDVESFTKVLKQKGLGGMVWALDDDDFAGFSCNGRYPLIQ 358
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
322 QQVPYAVKGNQWGYDDVKSMTKVLQKLNILGGAMISIDMDDTGKSCNGRYPLVQ 381
QY 359 TLQREL 364
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382 AVKRSI 387
RESULT 2
C3L1_HUMAN STANDARD; PRT; 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-39) (39 kDa synovial protein) (YKL-40).
GN CH13L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94064658; PubMed=8245017;
RA Hakala B.E.; White C.; Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family.";
RT J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97386591; PubMed=9244440;
RA Rehli M.; Krause S.W.; Andresen R.;
RT "Molecular characterization of the gene for human cartilage gp-39 (CH13L1), a member of the chitinase protein family and marker for late stages of macrophage differentiation.";
RT Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
```

```
RX MEDLINE=90328983; PubMed=2375755;
RA Nyirkos P.; Golds E.B.;
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine mammary protein expressed during the non-lactating period.";
RL Biochem. J. 269:265-268(1990).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG, PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
DR EMBL; M80927; AAA16074.1; -.
DR EMBL; Y08374; CAA69661.1; -.
DR EMBL; Y08375; CAA69661.1; JOINED.
DR EMBL; Y08376; CAA69661.1; JOINED.
DR EMBL; Y08377; CAA69661.1; JOINED.
DR EMBL; Y08378; CAA69661.1; JOINED.
DR PIR; S10677; S10677.
DR PIR; A33162; A33162.
DR Genew; HGNC:1932; CH13L1.
DR MIM; 601525; -.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD00471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Signal.
FT SIGNAL 1 21.
FT CHAIN 22 383 CHITINASE-3 LIKE PROTEIN 1.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 383 AA; 42613 MW; 76ADD829BEEC2D1 CRC64;
Query Match 43.4%; Score 1041.5; DB 1; Length 383;
Best Local Similarity 53.3%; Pred. No. 7.5e-72;
Matches 195; Conservative 66; Mismatches 96; Indels 9; Gaps 5;
QY 2 KLVCFYTNWAQYRQGEARFLPKDLPDLCTHLIYAFAGMTHNQLSTTEWDETLYQEFNG 61
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 KLVCFYTNWSQYREGDSCFPDLDRLFLCTHIYSFANISNDHIDTWENDVTLYGLMLT 82
QY 62 LKQWPKLTLIAIGWNGFTQKFTDMVATANNROTIVNSAIRFLKYSFDCGLDLDWEYP 121
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 LKRNPNLKTLLSVGWNFGSGFSGKIASNTQSRRTFTKSPVPPFLRTHGFGDLDAWLYP 142
QY 122 GSQSPAVDKERFTTLVQDLANAFQQAOTSGKERILLLSAAVPAQOTYVDAGYEVDKIAQ 181
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
143 GRR-----DKQHTTLIKEMKAEFIKEAQ-PGKQLLLSALSAGKVIDSYDAKISQ 196
QY 182 NLDVNLMAYPHGSWEKVTGNSPLYKROBESGAAASLNVDAAVQWLOKGTSPASKLIL 241
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
197 HLDFTISIMTYDFHGAWRGTGTHHSPLFRGQEDASPDREFTNTDYAVGYMLRGLAPASKLV 256
QY 242 GMPTVGRSFTLASSSDTRVGPATGSGTGPFTKEGGMLAYEVCWS-KGATKORIQOK 300
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
257 GIPTFGRSFTLA-SSETGVGAPISGPGIPGRFTKEAGTLAYEICDFLRGATVHRTLQOQ 315
QY 301 VPIYIFRDNQWGFDDVESFTKVLKQKGLGGMVWALDDDDFAGFSCNGQ-RYPLIOT 359
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
316 VPIYATKGNQWGYDDVKSQVQLKDRQLAGANVWALDDDDFAGFSCGQDLRPLTNA 375
QY 360 LRQELS 365
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```



121 PGSGSPAVDKERFTTLVODLANAFOEAOCTSGKERILLISAAVPAQTYVDAGYEVDKIA 180  
140 PGLRGSFARDRTWTFVLELLAQFNEAQLTWPRLLISAAVSGDPHVQKAYEARLLG 199  
181 QNLDFVNLMAIDFHGSWEKVTGHNPSLYKQEBESGAASLNDAVQOMLQKGTASKLI 240  
200 RLDDFISVLSDLHGSWEKVTGHNPSLPLGPKSSA-----YAMNYRQLGVPEKLL 254  
241 LGMPTVGRSFTLASSSDTRVGAATSGTGPPTKEGMLAYEYVCSW-KGATKQRIQOQ 299  
255 MGLPTVGRTHLLKASQNEAQLVGAATSGTGPPTKEGMLAYEYVCSW-KGATKQRIQOQ 314  
300 KVPYIFRDNQWGVDFDVSFETKVSYLKOKGLGAMVWALDDDDFAGFCSCNQGRIPLIOT 359  
315 YVPYAFKGEWGVYDDAISFGYKAFIKREHFGAMVWALDDDDFAGFCSCNQGRIPLIOT 374  
360 LRQEL 364  
375 LNNLL 379

RESULT 5  
OGP SHEEP  
ID OGP SHEEP STANDARD; PRT; 539 AA.  
AC Q28542; Q28543;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
DE (Oviductin) (Estrogen-dependent oviduct protein) (Estrus-associated  
DE oviductal glycoprotein) (OEGP).  
GN OVGPI OR OGP.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-39.  
RC TISSUE=Oviduct;  
RX MEDLINE=95269691; PubMed=7750470;  
RA Desouza M.M., Murray M.K.;  
RT "An estrogen-dependent secretory protein, which shares identity with  
RT chitinases, is expressed in a temporally and regionally specific  
RT manner in the sheep oviduct at the time of fertilization and embryo  
RT development.";  
RL Endocrinology 136:2485-2496(1995).  
RN [2]  
RP SEQUENCE OF 10-539 FROM N.A.  
RC STRAIN=Merino; TISSUE=Oviduct;  
RX MEDLINE=96329120; PubMed=8726871;  
RA Marshall J.T.A., Nancarrow C.D., Brownlee A.G.;  
RT "Cloning and sequencing of a cDNA encoding an ovine  
RT oestrus-associated oviductal protein.";  
RL Reprod. Fertil. Dev. 8:305-310(1996).  
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.  
CC -!- SUBCELLULAR LOCATION: Secretory granules.  
CC -!- TISSUE SPECIFICITY: OVIDUCT.  
CC -!- DEVELOPMENTAL STAGE: LEVELS ARE HIGHEST IN THE FIMBRIA AND AMPULLA  
CC AT ESTRUS AND ON DAY 1 OF PREGNANCY, WHEN GAMETE TRANSPORT AND  
CC FERTILIZATION OCCURS IN THE E2-DOMINATED FALLOPIAN TUBE. LEVELS  
CC DECLINE SIGNIFICANTLY ON DAY 2 AND UNDERGO A FURTHER SIGNIFICANT  
CC REDUCTION ON DAY 3 OF PREGNANCY COINCIDENT WITH TRANSPORT OF THE  
CC EMBRYO FROM THE OVIDUCT TO THE UTERUS, A REPRODUCTIVE STAGE  
CC ASSOCIATED WITH RISING PROGESTERONE LEVELS.  
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial  
entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; UI6719; AAC48471.1; -;  
DR EMBL; UI7988; AAB01052.1; -;  
DR InterPro; IPR001579; Chitinase\_18/2.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1;  
DR ProDom; PD000471; Glyco\_hydro\_18; 1;  
DR PROSITE; PS01095; CHITINASE\_18; FALSE\_NEG.  
KW Glycoprotein; Fertilization; Signal.  
FT SIGNAL 1 21 BY SIGNAL.  
FT CHAIN 22 539 OVIDUCT-SPECIFIC GLYCOPROTEIN.  
FT CARBOHYD 402 402 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 13 13 M -> V (IN REF. 2).  
FT CONFLICT 122 122 K -> N (IN REF. 2).  
FT CONFLICT 282 282 A -> V (IN REF. 2).  
FT CONFLICT 375 375 A -> V (IN REF. 2).  
FT CONFLICT 484 484 R -> H (IN REF. 2).  
FT CONFLICT 520 520 I -> T (IN REF. 2).  
SQ SEQUENCE 539 AA; 59535 MW; F35000269987C193 CRC64;  
Query Match 40.0%; Score 959.5; DB 1; Length 539;  
Best Local Similarity 51.0%; Pred. No. 2e-65;  
Matches 186; Conservative 57; Mismatches 115; Indels 7; Gaps 3;  
QY 2 KLVCFYTNWAOYRQGEARFLPKDLPDLCTHLIYAPAGTWHOLSTTEWDE-TLYQREN 60  
DB 23 KLVCFYTNWAFSRPGSASITLPRDLPFLCTHLVFAFASNNNQIVPKDLPDEKILYPEFN 82  
QY 61 GLKKNPKLTLAIGGWNFGTKETDMVATANNRQTFTVNSAIRFLKYSPDGLDLDWEY 120  
DB 83 KLKRNGLKTLISVGGWNFGTSRTKMLSTFSNERFVKSVIALLRTHGFGDGLDFLY 142  
QY 121 PGSGSPAVDKERFTTLVODLANAFOEAOCTSGKERILLISAAVPAQTYVDAGYEVDKIA 180  
DB 143 PGLRGSFARDRTWTFVLELLAQFNEAQLTWPRLLISAAVSGDPHVQKAYEARLLG 202  
QY 181 QNLDFVNLMAIDFHGSWEKVTGHNPSLYKQEBESGAASLNDAVQOMLQKGTASKLI 240  
DB 203 RLDDFISVLSDLHGSWEKVTGHNPSLPLGPKSSA-----YAMNYRQLGVPEKLL 257  
QY 241 LGMPTVGRSFTLASSSDTRVGAATSGTGPPTKEGMLAYEYVCSWKGATKQRIQOQ 299  
DB 258 MGLPTVGRTHLLKASQNEAQLVGAATSGTGPPTKEGMLAYEYVCSWKGATKQRIQOQ 317  
QY 300 KVPYIFRDNQWGVDFDVSFETKVSYLKOKGLGAMVWALDDDDFAGFCSCNQGRIPLIOT 359  
DB 318 YVPYAFKGEWGVYDDAISFGYKAFIKREHFGAMVWALDDDDFAGFCSCNQGRIPLIOT 377  
QY 360 LRQEL 364  
DB 378 LNNLL 382  
RESULT 6  
CHIT\_CABEL  
ID CHIT\_CABEL STANDARD; PRT; 617 AA.  
AC Q11174;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Putative endochitinase (EC 3.2.1.14).  
GN C04F6.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Nham M.;



Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-acetyl-D-glucosamine polymers of chitin.  
 -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL HYDROLASES).

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U42835; AAA83586.1; -;  
 WormPep; C04F6.3; CE03923.  
 InterPro; IPR002557; Chitin\_bind\_Pera.  
 InterPro; IPR001579; Chitinase\_18/2.  
 InterPro; IPR001223; Glyco\_hydro\_18.  
 Pfam; PF00704; Glyco\_hydro\_18; 1.  
 Pfam; PF01607; CBM 14; 2.  
 ProDom; PD000471; Glyco\_hydro\_18; 1.  
 SMART; SM00494; ChitB2; 2.  
 PROSITE; PS01095; CHITINASE\_18; 1.  
 Hypothetical protein; Hydrolase; Glycosidase; Chitin degradation.  
 ACT SITE 179 / 179 PROTON DONOR (BY SIMILARITY).  
 DOMAIN 435 / 478 THR-RICH.  
 SEQUENCE 617 AA; 56857 MW; DDA1D2AAAC0E54DA CRC64;

Query Match 39.9%; Score 956; DB 1; Length 617;  
 Best Local Similarity 40.8%; Pred. No. 4.5e-65;  
 Matches 195; Conservative 78; Mismatches 157; Indels 48; Gaps 10;

QY 5 CYFTNWAQYRGEARFLPKDLPSCILTHLYAFAGTNNHQLST-----TENWDEL 55  
 DB 57 CYFTNWAQYRGEARFLPKDLPSCILTHLYAFAGTNNHQLST-----TENWDEL 55  
 QY 56 YQEFNGLKKNPKLTLAIGGNFGTQKFTDMVATANNRQTFVNSAIRFLKYSDFGLD 115  
 DB 116 YRVNKLKVTDTQLKTLISFGGSGFTALFQGMASASRKVFIDSATFVTVGDFGID 175  
 QY 116 LDWEYVGSQSPVADKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQTYVDAGYE 175  
 DB 176 IDWEYV-----SGATDMANVVALVKELKAACEAGSTGKRLLVTAAGAATIDAGYD 231  
 QY 176 VDKIAQNLDVNLMAVDYFHGSWEKVTGHNPSLYKQEESSGAASLNVDAAVQWLOKCTP 235  
 DB 232 IPNLAPNFDILLMSYDFPGAWASLVGFNSPLYATTELPAAWGNVNDSSARYWQKGM 291  
 QY 236 ASKLIILGMPYGRSFTLASSDTRVGPATGSGTGPFT---KEGMLAYYEVCSW--KG 290  
 DB 292 KEKIIIVGMPYGRGWTNNASAINFGT---SGSPAKITQYVQAGVGAYFFCEMLANG 347  
 QY 291 ATKQRIQDKQVPIFRDNQWGFDDVESPKTKVSYLKQKGLGAMVWALDLPD-AGFSC 349  
 DB 348 ATRYWDSQVPLVQGNQWNSYDDESPANKWAVYKREGYGAFVWTLDFDDFNAGCSN 407  
 QY 350 NQGR-YPLIQLRQLSLPLP-----SGTPELEVPKGPQ-----S 385  
 DB 408 SNGQYPLISIVAKELGGVILPKKGVTTPATVTTVTGPPMTSAVTTTAAATTTT 467  
 QY 386 EPEHGPSQDQTFCCQKADGLPNRPRSSFFSCAAGRLFOQSCPTGLVFSNCKCCT 443  
 DB 468 RAATTITASNTNVCQSKSDGFPNSNCCGLFVLCLSSKSYNSCPSGLQYSASLKVCT 525

RESULT 7

OGP\_FIG

ID OGP\_FIG STANDARD; PRT; 527 AA.

AC Q28590;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
 DE (Oviductin) (Estrogen-dependent oviduct protein) (POSP-E3).  
 GN OVGPI OR OGP.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Yorkshire X Duroc X Hampshire; TISSUE=Oviduct;  
 RX MEDLINE=97107140; PubMed=8949888;  
 RA Bui W.C., Alvarez I.M., Choi I., Cleaver B.D., Simmen F.A.;  
 RT "Molecular cloning and characterization of an estrogen-dependent  
 RT porcine oviductal secretory glycoprotein.";  
 RL Biol. Reprod. 55:1305-1314(1996).  
 CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
 CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.  
 CC -!- SUBCELLULAR LOCATION: Secretory granules.  
 CC -!- TISSUE SPECIFICITY: OVIDUCT.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U43490; AAA85445.1; -;  
 InterPro; IPR001579; Chitinase\_18/2.  
 InterPro; IPR001223; Glyco\_hydro\_18.  
 Pfam; PF00704; Glyco\_hydro\_18; 1.  
 ProDom; PD000471; Glyco\_hydro\_18; 1.  
 PROSITE; PS01095; CHITINASE\_18; FALSE\_NEG.  
 KW Glycoprotein; Fertilization; Signal.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT CHAIN 22 527 OVIDUCT-SPECIFIC GLYCOPROTEIN.  
 FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 402 402 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 441 441 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 527 AA; 58519 MW; 31B78F49CA2363A2 CRC64;

Query Match 39.5%; Score 947.5; DB 1; Length 527;  
 Best Local Similarity 49.9%; Pred. No. 1.6e-64;  
 Matches 182; Conservative 60; Mismatches 116; Indels 7; Gaps 3;

QY 2 KLVCFYFTNWAQYRGEARFLPKDLPSCILTHLYAFAGTNNHQLSTTENWDEL-YQEPN 60  
 DB 23 KLVCFYFTNWAQYRGEARFLPKDLPSCILTHLYAFAGTNNHQLSTTENWDEL-YQEPN 82  
 QY 61 GLKKNPKLTLAIGGNFGTQKFTDMVATANNRQTFVNSAIRFLKYSDFGLDWEY 120  
 DB 83 QLKERNKELKTLISFGGSGFTALFQGMASASRKVFIDSATFVTVGDFGID 142  
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQTYVDAGYVDKIA 180  
 DB 143 PGLRSGPRDRWNFLLEELLAPREAQLTMRPRLLSAAVSADPHVIOKAYDVRLLG 202  
 QY 181 QNLDFVNLMAYDFHGSWEKVTGHNPSLYKQEESSGAASLNVDAAVQWLOKGTASKLI 240  
 DB 203 RLLDPIVLSYDLHGSWEKVTGHNPSLYKQEESSGAASLNVDAAVQWLOKGTASKLI 257  
 QY 241 LGMPYGRSFTLASSDTRVGPATGSGTGPFTKEGMLAYYEVCSWKGATKQ-IQDQ 299  
 DB 258 MGFTYGRSFTLASSDTRVGPATGSGTGPFTKEGMLAYYEVCSWKGATKQ-IQDQ 317  
 QY 300 KVPYIFRDNQWGFDDVESPKTKVSYLKQKGLGAMVWALDLPDFAFGSCNCRPLIQT 359  
 DB 318 YVPAYRGEWGVYDDESPANKWAVYKREGYGAFVWTLDFDDFNAGCSN 377  
 QY 360 LRQEL 364

```
Db 378 LNDLL 382
RESULT 8
ID OGP MOUSE STANDARD; PRT; 721 AA.
AC Q62010;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estragen-dependent oviduct protein)
GN OVGPI OR OGP OR CHITS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Oviduct;
RX MEDLINE=96115001; PubMed=7492680;
RA Sendai Y., Komiya H., Suzuki K., Onuma T., Kikuchi M., Hoshi H.,
Araki Y.;
RT "Molecular cloning and characterization of a mouse oviduct-specific
glycoprotein.";
RL Biol. Reprod. 53:285-294(1995).
CC -!- FUNCTION: BINDS TO OCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF THE OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC [1]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D32137; BAA06863.1; -.
CC MGD; MGI:106661; Ovgpl.
CC InterPro; IPR001579; Chitinase_18/2.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21
FT CHAIN 22 721
FT DOMAIN 486 632
FT [TAP]-G-[IV]
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 721 AA; 78807 MW; 37246C8F01665652 CRC64;
Query Match 39.3%; Score 942; DB 1; Length 721;
Best Local Similarity 49.3%; Pred. No. 6.5e-64;
Matches 185; Conservative 60; Mismatches 122; Indels 8; Gaps 4;
QY 2 KLVCFYTNNAQYRQGEARFLPKDPLSLCTHLYAFAGMTNHSLSLSTEWNDE-TLYQEEN 60
Db 23 KLVCFYTNNAHSRPGASIMPHDLPFLCTHLIFAFASMSNNQIVAKNQDENLVPEEN 82
QY 61 GLKKNPKLTLIAIGGNWFGTKFTDMVATANNROTFFNSAIRFKYKVSFGDLIDWEY 120
Db 83 KLKERNRELKTLISIGGNWFGTSRFTAMLSLANREKFDSDVSIFRIRHGFDGLDFFLY 142
QY 121 PGSQSPAVDKERFTTLVDLANAFOEQAQTSCKERILLISAAVPAQGTVDYAGYVDKIA 180
Db 143 PGLRSGPPHWRNLFLLIBELQAFERALLQHPRLLSAAVSGIPSIHTSYDALLG 202
```

```
QY 181 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKROBESGAASLNVDAAVQOVLKGTASKLI 240
Db 203 RRLDFINVLSDYDLHGSWEKFTGHNPLFLSLPDSKSSA-----YAMNYWRKLGTPADKLI 257
QY 241 LGMPTYGSRFTLASSSDTRVGAPATGSGTGPFTKEGGMLAYEYVCSW-KGATKQRIQDO 299
Db 258 MGFPTYGRNFYLLKESKNGLOQTASWGPASPGKYTKQAGFLAYEYVCSFVQRAKRWIDYQ 317
QY 300 KVPYIFRDNOWGVFDDVSEFKTKVSYLKQKGLGGAMVWALDLDLDFAGSCNCRYPLOT 359
Db 318 YVPYAFKGEWLGYDDTISFSYKAMVYVREHFGGANVWTLDMDDVYRGTCFNGNGPFLVHI 377
QY 360 LRQELSLPYLPSGTP 374
Db 378 L-NEILLVQTESNSTP 391
RESULT 9
ID OGP MESAU STANDARD; PRT; 671 AA.
AC Q60557; Q60526;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estragen-dependent oviduct protein) (ZP-0).
GN OVGPI OR OGP.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Oviduct;
RX MEDLINE=96115007; PubMed=7492686;
RA Suzuki K., Sendai Y., Onuma T., Hoshi H., Hiroi M., Araki Y.;
RT "Molecular characterization of a hamster oviduct-specific
glycoprotein.";
RL Biol. Reprod. 53:345-354(1995).
RN [2]
SEQUENCE OF 14-671 FROM N.A., AND REVISIONS.
RP TISSUE=Oviduct;
RC TISSUE=Oviduct;
RA Paquette Y.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE OF 22-39.
RX MEDLINE=94058981; PubMed=8240241;
RA Malette B., Bleau G.;
RT "Biochemical characterization of hamster oviductin as a sulphated
zona pellucida-binding glycoprotein.";
RL Biochem. J. 295:437-445(1993).
CC -!- FUNCTION: BINDS TO OCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC MIGHT ACT AS A PROTECTIVE SECRETION INFLUENCING THE FIRST STEPS OF
THE REPRODUCTIVE PROCESS NECESSARY FOR THE NORMAL TRIGGERING OF
FERTILIZATION AND EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- PTM: HIGHLY O-GLYCOSYLATED AND ALSO N-GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL; D32218; BAA06977.1; -  
EMBL; U15048; AAC3584.1; -  
InterPro; IPR001579; Chitinase 18/2.  
InterPro; IPR001223; Glyco\_hydro\_18.  
Pfam; PF00704; Glyco\_hydro\_18; 1.  
ProDom; PD000471; Glyco\_hydro\_18; 1.  
PROSITE; PS01095; CHITINASE 18; FALSE NEG.  
KW Glycoprotein; Fertilization; Repeat; Signal.  
FT SIGNAL 1 21 OVIDUCT-SPECIFIC GLYCOPROTEIN.  
FT CHAIN 22 671 8 X 15 AA TANDEM REPEATS.  
FT DOMAIN 490 609 1.  
FT REPEAT 490 504 2.  
FT REPEAT 505 519 3.  
FT REPEAT 520 534 4.  
FT REPEAT 535 549 5.  
FT REPEAT 550 564 6.  
FT REPEAT 565 579 7.  
FT REPEAT 580 594 8.  
FT REPEAT 595 609 9.  
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 511 511 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 526 526 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 541 541 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 596 596 C -> A (IN REF. 4).  
FT CONFLICT 26 26 H -> I (IN REF. 4).  
FT CONFLICT 33 33 D -> G (IN REF. 2 AND 3).  
FT CONFLICT 137 137 R -> Q (IN REF. 2 AND 3).  
FT CONFLICT 153 153 F -> Y (IN REF. 2 AND 3).  
FT CONFLICT 165 165 Q -> L (IN REF. 2 AND 3).  
FT CONFLICT 193 193 MISSING (IN REF. 2 AND 3).  
FT CONFLICT 531 545 T -> I (IN REF. 2 AND 3).  
FT CONFLICT 595 595  
FT SEQUENCE 671 AA; 73250 MW; BB57E0E514EC1972 CRC64;

Query Match 38.7%; Score 927; DB 1; Length 671;

Best Local Similarity 46.9%; Pred. No. 8.1e-63;

Matches 184; Conservative 61; Mismatches 135; Indels 12; Gaps 4;

QY 2 KLVCFYFTWNAOYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNHQLSTTEWDE-TLYQEFN 60  
DB 23 KLVCFYFTWNAHSRPPVASTLPRLDLPFLCTHLIFAFASNNQIVANNLQDEKILYPEFN 82  
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFTVNSAIRFLKYSFDGLDWEY 120  
DB 83 KLKERNRALKTLISVGGWNGFTRTMTLSAREKFTIGSVSLRTHGFGLDLDFLY 142  
QY 121 PGSQSPAVDKERFTTLVODLANAQFQAQTSKGRLLLSAAVPAQGTVDYAGYEVKIA 180  
DB 143 PGLRGSPINDRNWFLFELQFAPEKALLTQRPRLLSAAVSGIPYLIQTSYDVHLLG 202  
QY 181 QNLDFVNLWAYDFHGSWEKVTCHNSPLKYRQESGAASLNVDAVQWLQGTFTASKLI 240  
DB 203 RRLDPTINVLSDYHGSWEKSTGHNSPLFSLPDPKSSA-----FAMNYWRNLGAPADKLL 257  
QY 241 LGMPYGRSFTLASSSDTRVRGAPATGSGTPGFTKEGGMLAYEYVCSW-KGATKQRIQDQ 299  
DB 258 MGFPAVGRFTLLRRESKNGLQASNGPASPFGYTKQAGFLAYEYVCSFIQRAEKWIHQ 317  
QY 300 KVPYIFRDNQWGFDDVSFKTKVSYLKQKGLGGAMWALDLDLDDPAGSCNGRYPLIQT 359  
DB 318 YVPYAYKGEWGYDDAVSFSYKAMFVKHFGGAMVWTLDDMDVRGTFCGNGPPLVHI 377  
QY 360 L-----RQELSLPYLPSGTPELEVKPGQPS 386  
DB 378 LNELLVRAEFNSTPLPQFMFTLPLVNSSGPGSE 409

RESULT 10  
OGP PAPAN

ID\_ OGP PAPAN STANDARD; PRT; 623 AA.

AC P36718;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
DE (Oviductin) (Estrogen-dependent oviduct protein).  
GN OVEP1 OR OGP.  
OS Papio anubis (Olive baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Papio.  
OX NCBI\_TaxID=9555;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Oviduct;  
RX MEDLINE=98244335; PubMed=9584944;  
RA Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,  
RA Donnelly K.M., Arias E.B., Jaffe R.C.;  
RT "The baboon oviduct: characteristics of an oestradiol-dependent  
RT oviduct-specific glycoprotein.";  
RL Hum. Reprod. Update 3:541-552(1997).  
RN [2]  
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
RC TISSUE=Oviduct;  
RX MEDLINE=91367180; PubMed=1716345;  
RA Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,  
RA Jaffe R.C.;  
RT "Cloning of a recombinant complementary DNA to a baboon (Papio  
RT anubis) estradiol-dependent oviduct-specific glycoprotein.";  
RL Mol. Endocrinol. 5:356-364(1991).  
CC -1- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.  
CC -1- SUBCELLULAR LOCATION: Secretory granules.  
CC -1- TISSUE SPECIFICITY: OVIDUCT.  
CC -1- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.  
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL; M59903; AAB39765.1; -  
PIR; A37954; A37954.

InterPro; IPR001579; Chitinase 18/2.

InterPro; IPR001223; Glyco\_hydro\_18.

Pfam; PF00704; Glyco\_hydro\_18; 1.

ProDom; PD000471; Glyco\_hydro\_18; 1.

PROSITE; PS01095; CHITINASE 18; FALSE\_NEG.

KW Glycoprotein; Fertilization; Signal.

FT SIGNAL 1 21 BY SIMILARITY.

FT CHAIN 22 623 OVIDUCT-SPECIFIC GLYCOPROTEIN.

FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).

FT SEQUENCE 623 AA; 69291 MW; 9B21CB481FFF1268 CRC64;

Query Match 38.6%; Score 925.5; DB 1; Length 623;

Best Local Similarity 44.1%; Pred. No. 9.6e-63;

Matches 194; Conservative 66; Mismatches 135; Indels 45; Gaps 8;

QY 2 KLVCFYFTWNAOYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNHQLSTTEWDE-TLYQEFN 60  
DB 23 KLVCFYFTWNAHSRPPVASTLPRLDLPFLCTHLIFAFASNNQIVANNLQDEKILYPEFN 82  
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFTVNSAIRFLKYSFDGLDWEY 120  
DB 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFTVNSAIRFLKYSFDGLDWEY 120

```

Db 83 KLKERNRELKTLISIGGNWFGTSRFTTLMSTFANREKFIASVISLRLTHDFDGLDLFFLY 142
QY 121 PGSQSPAVDKERFTTLVODLANAFOEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 180
Db 143 PGLRGSPMDHRTWFLFIEELLFAFKREALTWPRLLLSAAVGVPHIVQTSYDVRFLG 202
QY 181 QNLDFVNLMAYPHGSWEKVTGHNSPLYKQESGAAALNVDAAVQOVLQKGTTPASKLI 240
Db 203 RLDDFINVLSYDLHGSWEKFTGHNSPLSLPEDPKSSA-----YAMNYWRKLGAPSEKLI 257
QY 241 LGMPTVGRSFTLASSSDTRVGAPATSGTGPFTKEGMLAYEVCSS--WKGATKQRIOD 298
Db 258 MGITYGRTFRLKASKNGLOAIGPASPGRKYTKQAGFLAYFEICSFVW-GAKKHWDY 316
QY 299 OKVPYIFRDNQWGFDDVESFKTKVSLKQGLGGAMVWALDLDPAFSCNCGRYPLIQ 358
Db 317 QYVPYANKGEWGYDDAISFSYKAWFIRKHFGGAMVWTLDDVDRGFTCGTGPFLVY 376
QY 359 TL-----RQELSLPYLP-----SGTPELEVPKPGQSPSEPHGP 391
Db 377 VMNDILVRAEFSSTSLPQFWLSSAVNSSSTDPERLAVTKAWTTDIKILPG-----GEAGV 432
QY 392 SPQDFTFCQKADGLPNR 411
Db 433 -----TEHKGKCNMTITPR 447

RESULT 11
OGP_HUMAN STANDARD; PRT; 678 AA.
AC Q12889; Q15841;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein).
GN OVGPI OR OGP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=95119256; PubMed=7819450;
RA Arias E.B., Verhage H.G., Jaffe R.C.;
RT "Complementary deoxyribonucleic acid cloning and molecular
RT characterization of an estrogen-dependent human oviductal
RT glycoprotein."
RL Biol. Reprod. 51:685-694 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Jaffe R.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U09550; AAA86946.1; -
DR EMBL; U58010; AAB04126.1; -
DR EMBL; U58001; AAB04126.1; JOINED.
DR EMBL; U58002; AAB04126.1; JOINED.
DR EMBL; U58003; AAB04126.1; JOINED.

```

```

DR EMBL; U58004; AAB04126.1; JOINED.
DR EMBL; U58005; AAB04126.1; JOINED.
DR EMBL; U58006; AAB04126.1; JOINED.
DR EMBL; U58007; AAB04126.1; JOINED.
DR EMBL; U58008; AAB04126.1; JOINED.
DR EMBL; U58009; AAB04126.1; JOINED.
DR Genew; HGNC:8524; OVGPI.
DR MIM; 603578; -
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE 18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21
FT CHAIN 22 678 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 477 477 M -> T (IN REF. 2).
FT CONFLICT 511 511 S -> P (IN REF. 2).
FT CONFLICT 514 514 Y -> H (IN REF. 2).
FT CONFLICT 676 676 E -> Q (IN REF. 2).
SQ SEQUENCE 678 AA; 75421 MW; 245F2CEDCE92768B CRC64;

Query Match 38.3%; Score 917.5; DB 1; Length 678;
Best Local Similarity 44.8%; Pred. No. 4.4e-62;
Matches 193; Conservative 66; Mismatches 145; Indels 27; Gaps 7;

QY 2 KLVCYFTNWAQVROGEARFLPKDLPSLCETHLIYAFAGMTNHLSTTWNDE-TLYQEFN 60
Db 23 KLVCYFTNWAHSRPGPASTLPHLDPLCLTHLIYAFASNNNNQIVAKDLQDEKILYPERN 82
QY 61 GLKWNPKLTLAIGGNWFGTKFTDMVATANNQTFVNSAIRFLKYKSFGLDLWBY 120
Db 83 KLKERNRELKTLISIGGNWFGTSRFTTLMSTFANREKFIASVISLRLTHDFDGLDLFFLY 142
QY 121 PGSQSPAVDKERFTTLVODLANAFOEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 180
Db 143 PGLRGSPMDHRTWFLFIEELLFAFKREALTWPRLLLSAAVGVPHIVQTSYDVRFLG 202
QY 181 QNLDFVNLMAYPHGSWEKVTGHNSPLYKQESGAAALNVDAAVQOVLQKGTTPASKLI 240
Db 203 RLDDFINVLSYDLHGSWEKFTGHNSPLSLPEDPKSSA-----YAMNYWRKLGAPSEKLI 257
QY 241 LGMPTVGRSFTLASSSDTRVGAPATSGTGPFTKEGMLAYEVCSS--WKGATKQRIOD 298
Db 258 MGITYGRTFRLKASKNGLOAIGPASPGRKYTKQAGFLAYFEICSFVW-GAKKHWDY 316
QY 299 OKVPYIFRDNQWGFDDVESFKTKVSLKQGLGGAMVWALDLDPAFSCNCGRYPLIQ 358
Db 317 QYVPYANKGEWGYDDAISFSYKAWFIRKHFGGAMVWTLDDVDRGFTCGTGPFLVY 376
QY 359 TL-----RQELSLPYLP-----SGTPELEVPKPGQSPSEPHGPSPQCD--TFQC 400
Db 377 VLNDILVRAEFSSTSLPQFWLSSAVNSSSTDPERLAVTKAWTTDTSKILPPGGEAGVTEH 436
QY 401 GKADGLYNPR 411
Db 437 GKCNMTITPR 447

RESULT 12
CHIT_BRUMA STANDARD; PRT; 504 AA.
ID CHIT BRUMA
AC P29030;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Endochitinase precursor (EC 3.2.1.14) (MFI antigen).
OS Brugia malayi.

```

CC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
OX Onchocercidae; Brugia.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=92179220; PubMed=1542646;  
RA Fuhrman J.A., Lane W.S., Smith R.F., Piesens W.F., Perler F.B.;  
RT "Transmission-blocking antibodies recognize microfilarial chitinase  
in brugian lymphatic filariasis."  
RL Proc. Natl. Acad. Sci. U.S.A. 89:1548-1552(1992).  
CC -!- FUNCTION: THE MFI ANTIGEN IS A MICROFILARIAL CHITINASE, WHICH MAY  
FUNCTION TO DEGRADE CHITIN-CONTAINING STRUCTURES IN THE MICRO-  
FILARIA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND  
TRANSMISSION.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
acetyl-D-glucosamine polymers of chitin.  
CC -!- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MFI ANTIGEN CORRESPOND  
WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.  
CC -!- PTM: O-GLYCOSYLATED.  
CC -!- MISCELLANEOUS: KNOWN TO BIND CALCIUM.  
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M73689; AAA27854.1; -;  
DR PIR; A38221; A38221.  
DR InterPro; IPR002557; Chitin\_bind\_Pera.  
DR InterPro; IPR001579; Chitinase\_18/2.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR Pfam; PF01607; CBM\_14; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR SMART; SM00494; ChtBD2; 1.  
DR PROSITE; PS01095; CHITINASE\_18; 1.  
DR Repeat; Glycosidase; Chitin degradation; Signal; Antigen;  
KW Repeat; Glycoprotein; Calcium-binding.  
FT SIGNAL 1 22  
FT CHAIN 23 504 ENDOCHITINASE.  
FT DOMAIN 23 400 CATALYTIC.  
FT DOMAIN 401 450 SER/THR-RICH (LINKER).  
FT DOMAIN 407 448 3 X 14 AA APPROXIMATE TANDEM REPEATS.  
FT ACT SITE 148 / 148 PROTON DONOR (BY SIMILARITY).  
SQ SEQUENCE 504 AA; 55971 MW; A78BE7BF8BE3709B CRC64;  
Query Match 35.9%; Score 860.5; DB 1; Length 504;  
Best Local Similarity 37.0%; Pred. No. 6.2e-58;  
Matches 177; Conservative 80; Mismatches 170; Indels 51; Gaps 9;  
5 CYFTNWAQYRQGEARFLPKDLPSCLTILYAFAGMTHQLSTT---EWNDST-----L 55  
27 CYFTNWAQYRDEGEKFLGPNGLCTILYAFKV--DELGDSKPFENWEDSTWSKM 84  
56 YQEFNGLKKNPKLKTLLAIGGNFGTQKTDWVATANNRQTTFVNSAIRFLKYSFDGLD 115  
85 YSAVTKLRNTPGLKVLISYCGYNGFSIAFTGIAXSAQKTERFIKSAIAFLKNNFDGFD 144  
116 LDWEYPGSGSPAVDKERTTTLVQDLANAFQEAQTSKGERILLNAAVPAQTYVDAGYE 175  
145 LDWEYFVGVA-----EEHAKLVAMKTAFAVEAKTSKQRLLLTAASVAGKTTDGSYN 198  
176 VDKIAQNLDVNLMAIDYFHGSKEVTVGHSPLYKQESGAASLNVDAAVQVQLQKQTFP 235  
199 VESLGNFDLLFMSYDLHGSKEKVDLHCKLHPKTEGVSGIGFNTPEAAYWASKGMP 258  
236 ASKILGMPTYGRSTFLASSSDSTRVCAPATGSGTFCPFTKEGGMLAYVEVCSW-KGATKQ 294

Db 259 KEKIIIGIMYAQQWTLNPNSETAIGAAASRSPSSASKTNPAGGTAYWEICKYLKEGGKE 318  
Qy 295 RIQOKV--PYIFRDNQWGFDDVESFKTKVSLYKOKGLGGAMVWALDLDLDDPAGESCNOGR 353  
Db 319 TVHQGVGAYWVGQWGYDNEETIRIKMKWKEGYGGAFIWDLDFDDFTGKSCGKGP 378  
Qy 354 YPLIQTLORELSLPLPGSTPELEVPKPG-----QPSE 386  
Db 379 YPLNNAISSELE---GESENPEITTEPSITETEAYETDETEETSETAYDTDEETSE 435  
Qy 387 PEHGPSQDFTFC--GXADGLYNPRRSFYSACAGRLFQOOSCTGLVFSNSCKCC 442  
Db 436 TEATTYDTDETEGQECRPERDGLFPHPTDCHLFIQCANNIAYVMQCPATFFFNDAIKVC 493  
RESULT 13  
CHIT MANSE  
ID CHIT MANSE STANDARD; PRT; 554 AA.  
AC P36362;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Endochitinase precursor (EC 3.2.1.14).  
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;  
OC Ditrysia; Sphingioidea; Sphingidae; Sphinginae; Manduca.  
OX NCBI\_TaxID=7130;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93357793; PubMed=8353525;  
RA Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.;  
RT "Sequence of a cDNA and expression of the gene encoding epidermal and  
gut chitinases of Manduca sexta."  
RL Insect Biochem. Mol. Biol. 23:691-701(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97215580; PubMed=9061927;  
RA Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.;  
RT "Isolation and characterization of a genomic clone for the gene of an  
insect molting enzyme, chitinase."  
RL Insect Biochem. Mol. Biol. 27:37-47(1997).  
CC -!- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING  
PROCESS.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
acetyl-D-glucosamine polymers of chitin.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: EPIDERMIS AND GUT.  
CC -!- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,  
BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH  
INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A  
RAPID DECLINE IS SEEN. IN THE GUT IS DETECTED ON DAY 6 WITH LOWER  
LEVELS SEEN ON DAYS 0, 7 AND 8.  
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U02270; AAC04924.1; -;  
DR EMBL; L49234; AAB53952.1; -;  
DR InterPro; IPR002557; Chitin\_bind\_Pera.  
DR InterPro; IPR001579; Chitinase\_18/2.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR Pfam; PF01607; CBM\_14; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR SMART; SM00494; ChtBD2; 1.

DR PROSITE; PS01095; CHITINASE 18; 1.  
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein.  
FT SIGNAL 1 39 POTENTIAL.  
FT CHAIN 20 554 ENDOCHITINASE.  
FT DOMAIN 396 453 SER/THR-RICH.  
FT ACT SITE 146 146 PROTON DONOR (BY SIMILARITY).  
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 554 AA; 62203 MW; 39890756C96CD490 CRC64;  
  
Query Match 34.4%; Score 824; DB 1; Length 554;  
Best Local Similarity 33.7%; Pred. No. 4.2e-55;  
Matches 181; Conservative 87; Mismatches 159; Indels 110; Gaps 15;  
  
QY 1 AKLVCFYTNWAOYRQGEARFLPKDLPSCETHLIYAFAGTNNHQLSTTWNDETLL----- 55  
DB 23 ARIVCYFSNVAVYRPGVGRYGIEDIPEVKCTHIYFVIGV-----TEGNSVLLIDPE 75  
QY 56 -----YQFENGLKQWPKLTLALGAGWNETQKFTDMVATANNROTQVNSAIRFLPK 108  
DB 76 LDVDKNGFRNFTSLRSHSPSVKPMVAVGWAEGSSKSHVAKQSTMRFSIRVSVFLKK 135  
QY 109 YSPDGLDLWEYPGS--QGSPAVDKERFTTLVODLANAQOEAQTSKGERLLLSAAVPAG 166  
DB 136 YQFDGLDLWEYFGAADRGGSDKDKFLYLQELRRAIRVCK--GWE---LTAAPLA 190  
QY 167 QTVVDAGYEVDKTAQNLDFNLWAYDFHGSWEKVTGHNISPLYKROESGAASLNDAV 226  
DB 191 NFRLEMGYHVPCLQELDAIHVMSYDLRGWAGFADVHSPLYKRPDQWAYEKLNVNDGL 250  
QY 227 QQWLQGTASKLILGMPYGRSFTLASSD-----TRVCAPATGSGTGPPTKEGMLA 281  
DB 251 HLWEKQPSNKLWGPYIPYRFTLSAGNNVYGLTFINKEA--GGDPAFYNTATGFWA 309  
QY 282 YVEVCSW-----KGATKQRIQDKQVYIFRDNQWGFDDVESKTVSYLKQKGLGAMV 336  
DB 310 YVEICTEVDKDSGWTKWDEQCKPYAYKGTQWGVYEDPSRVEIKWNIKQKYLGMAT 369  
QY 337 WALDLDFAGFCNQRYPILQTLRELSLYLP-----SGTPELVKPKQPSPEH 389  
DB 370 WAIMDDDFQGL--CGE--KNPLIKILHKHMSYVPPPHPTNTTTPTEWARP--PSTPSDPSE 426  
QY 390 G-----PSPQDT--- 397  
DB 427 GDIPTTTAKPASTTKVTKTTTTTAKPQSVDEENDINVRPEKPEPEPEVEVP 486  
QY 398 -----FCQKADGLYPNPRSSFSYCAAGRLFOQSCPTGLVFSNCKCCTW 444  
DB 487 PTENEVDGSEICNSDDYI--PDKKHCDKYRCVNGEAMQFSCQHGTVFVNLNVCDM 542  
  
RESULT 14  
CH11\_BACCI  
ID\_CH11\_BACCI STANDARD; PRT; 699 AA.  
AC P2053;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Chitinase A1 precursor (EC 3.2.1.14).  
GN CH1A1.  
OS Bacillus circulans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1397;  
RN [1]  
RC STRAIN=WL-12;  
RX MEDLINE=90368776; PubMed=2203782;  
RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;  
RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed  
its evolutionary relationship to Serratia chitinase and to the type  
III homology units of fibronectin.";

J. Biol. Chem. 265:15659-15665 (1990).  
[2]  
RN MUTAGENESIS.  
RC STRAIN=WL-12;  
RX MEDLINE=93366760; PubMed=8103047;  
RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,  
Uchida M., Tanaka H.;  
RT "Identification of glutamic acid 204 and aspartic acid 200 in  
chitinase A1 of Bacillus circulans WL-12 as essential residues for  
chitinase activity";  
RL J. Biol. Chem. 268:18567-18572 (1993).  
CC -|- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
acetyl-D-glucosamine polymers of chitin.  
CC -|- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC -|- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL; M57601; AAA81528.1; .  
DR PIR; A38368; A38368.  
DR HSSP; P02751; LFNA.  
DR InterPro; IPR003610; CEM\_5\_12.  
DR InterPro; IPR001579; Chitinase\_18/2.  
DR InterPro; IPR003961; FN.III.  
DR InterPro; IPR003962; FN.II repeat.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00041; fn3; 2.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR Pfam; PF02839; CEM\_5\_12; 1.  
DR PRINTS; PRO0014; FNTYPE.III.  
DR ProDom; PD000471; Glyco\_hydro\_18; 2.  
DR SMART; SM000495; ChtBD3; 1.  
DR SMART; SM00060; FN3; 2.  
DR PROSITE; PS01095; CHITINASE 18; 1.  
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Repeat.  
FT SIGNAL 1 41  
FT CHAIN 42 699 CHITINASE A1.  
FT DOMAIN 42 460 CATALYTIC.  
FT DOMAIN 465 549 FIBRONECTIN TYPE-III (R-1).  
FT DOMAIN 560 644 FIBRONECTIN TYPE-III (R-2).  
FT ACT SITE 204 204 PROTON DONOR (PROBABLE).  
FT MUTAGEN 200 200 D->N: DECREASE IN ACTIVITY.  
FT MUTAGEN 200 200 D->E: NO CHANGE IN ACTIVITY.  
FT MUTAGEN 204 204 E->D,Q: LOSS OF ACTIVITY.  
SQ SEQUENCE 699 AA; 73677 MW; AC7C9B2E2E987643 CRC64;  
  
Query Match 24.4%; Score 584.5; DB 1; Length 699;  
Best Local Similarity 33.0%; Pred. No. 9.3e-37;  
Matches 147; Conservative 71; Mismatches 143; Indels 85; Gaps 16;  
  
QY 2 KLVCYFTWAOYRQGEARFLPKDLPSCETHLIYAFAGT----- 40  
DB 45 KIVGYFSAWYAGR---NYNVADIDPTKVTHTNYAFADICWNGIHGNDPPSPNPTWTC 101  
QY 41 TNHQLSTTE-----WMD-----ETLYQ-----EFNGLKQWPKLTLAI 75  
DB 102 QNEKQTNVNPNGTIVLGDPPWIDTCKTFAGDTWDQPIAGNINQLNKLQTNENLKIISV 161  
QY 76 GGNFQTKPTDMVATANNROTQVNSAIRFLPKYFSDGLDLWEYPGSOG-----SPAVD 130  
DB 162 GGWTW--SNRFSVAAATAATREVFANSVAVDFLRKYNFDGVDLDWEYFVSGGLDGNKRPED 220  
QY 131 KERFTTLVODLANAQOEAQTSKGERLLLSAAVPAGQTVVDAGYVDKTAQNLDFNLMA 190  
DB 221 KQNTILLISKIREKLDAGAVDGK--KYLTTIASGASATYA--ANTELAKIAIVDWININT 278



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:55:24 ; Search time 25.0148 Seconds  
(without alignments)  
3665.470 Million cell updates/sec

Title: US-10-004-219b-10

Perfect score: 2398

Sequence: 1 AKLVCFYTNWQYRQGEARF.....QQSCPTGLVFSNCKRCTWN 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virius.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	2398	100.0	466	4	Q13231		Q13231 homo sapien
2	1935	80.7	387	4	Q9H3V8		Q9H3V8 homo sapien
3	1334	55.6	396	11	Q9D7Q1		Q9D7Q1 mus musculus
4	1323	55.2	476	4	Q9B2P6		Q9B2P6 homo sapien
5	1309.5	54.6	472	11	Q9JLN1		Q9JLN1 mus musculus
6	1309.5	54.6	473	11	Q9D803		Q9D803 mus musculus
7	1304	54.4	472	6	Q95M17		Q95M17 bos taurus
8	1302.5	54.3	488	13	Q90W34		Q90W34 bufo japoni
9	1301.5	54.3	473	11	Q99PH2		Q99PH2 mus musculus
10	1053.5	43.9	352	11	Q9WTV1		Q9WTV1 rattus norv
11	1046.5	43.6	383	4	Q96HT7		Q96HT7 homo sapien
12	1041.5	43.4	380	4	Q96F97		Q96F97 homo sapien
13	1029.5	42.9	383	6	Q8SFQ0		Q8SFQ0 capra hircu
14	1021.5	42.6	383	6	Q29411		Q29411 sus scrofa
15	1015.5	42.3	696	5	Q9V2V2		Q9V2V2 drosophila
16	1015.5	42.3	1013	5	Q960M0		Q960M0 drosophila

17	991.5	41.3	381	11	Q99J84	Q99J84 mus musculus
18	991	41.3	368	4	Q9ULY4	Q9ULY4 homo sapien
19	979	40.8	332	6	O18949	O18949 bos taurus
20	975.5	40.7	365	11	Q91X89	Q91X89 mus musculus
21	970.5	40.5	398	11	Q91Z98	Q91Z98 mus musculus
22	970	40.5	402	11	Q8VH01	Q8VH01 mus musculus
23	968.5	40.4	398	11	P70201	P70201 mus musculus
24	967.5	40.3	398	11	O35744	O35744 mus musculus
25	965	40.2	402	11	O8VH43	O8VH43 mus musculus
26	926.5	38.6	624	6	O19118	O19118 macaca mula
27	919.5	38.3	4498	5	Q9W2Z3	Q9W2Z3 drosophila
28	883	36.8	525	5	O44079	O44079 anopheles g
29	868.5	36.2	504	5	Q9NGK8	Q9NGK8 wuchereria
30	864.5	36.1	457	6	Q95LB3	Q95LB3 oryctolagus
31	864.5	36.1	483	5	O23737	O23737 chelonus sp
32	858	35.8	520	5	O17100	O17100 acanthochei
33	856	35.7	497	5	O25615	O25615 onchocerca
34	855	35.7	524	5	O27919	O27919 acanthochei
35	848	35.4	460	5	Q9W2M7	Q9W2M7 drosophila
36	827.5	34.5	553	5	P91731	P91731 hyphantria
37	823	34.3	543	5	Q9GR93	Q9GR93 bombyx mori
38	822	34.3	566	5	O8WR52	O8WR52 bombyx mori
39	821	34.2	552	5	Q9GV44	Q9GV44 spodoptera
40	820	34.2	543	5	Q9GV05	Q9GV05 bombyx mori
41	820	34.2	565	5	P90710	P90710 bombyx mori
42	818	34.1	544	5	Q9GQC4	Q9GQC4 bombyx mori
43	816.5	34.0	467	5	O15993	O15993 penaeus jap
44	814	33.9	565	5	Q9GPG9	Q9GPG9 bombyx mand
45	810	33.8	460	5	Q8WS95	Q8WS95 glossina mo

#### ALIGNMENTS

RESULT 1

Q13231 ID Q13231 PRELIMINARY; PRT; 466 AA.

AC Q13231; DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)  
DE Chitotriosidase precursor.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1].  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96064695; PubMed=7592832;  
RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;  
RT "Cloning of a cDNA encoding chitotriosidase, a human chitinase  
RT produced by macrophages";  
RL J. Biol. Chem. 270:26252-26256 (1995).  
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
DR EMBL; U29615; AAC50246.1; -;  
DR InterPro: IPR001579; Chitinase\_18/2.  
DR InterPro: IPR002557; Chitin\_bind\_Pera.  
DR InterPro: IPR001223; Glyco\_hydro\_18.  
DR Pfam: PF01607; CBM\_14; 1.  
DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
DR ProDom: PD000471; Glyco\_hydro\_18; 1.  
DR SMART: SM00494; ChitBD2; 1.  
DR PROSITE: PS01095; CHITINASE\_18; 1.  
KW Glycosidase; Hydrolase; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 466 CHITOTRIOSIDASE.  
SQ SEQUENCE 466 AA; 51681 MW; B4312D1E885E386D CRC64;

Query Match 100.0%; Score 2398; DB 4; Length 466;  
Best Local Similarity 100.0%; Pred. No. 1.1e-177;  
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHLSTTEWNETLYQEFN 60  
 DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHLSTTEWNETLYQEFN 81  
 QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEY 120  
 DB 82 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEY 141  
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAVPAQOTYVDAGYEDVKIA 180  
 DB 142 PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAVPAQOTYVDAGYEDVKIA 201  
 QY 181 QNLDFVNLMAVDPHGSWEKVTGHSNPLKYRQESGAAASLNVDAAVQOVLQKGTTPASKLI 240  
 DB 202 QNLDFVNLMAVDPHGSWEKVTGHSNPLKYRQESGAAASLNVDAAVQOVLQKGTTPASKLI 261  
 QY 241 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGMLAYEVCWSKGGATKQRIQDQK 300  
 DB 262 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGMLAYEVCWSKGGATKQRIQDQK 321  
 QY 301 VPYIFRDQNWGVDDVESFPTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRIPIQTL 360  
 DB 322 VPYIFRDQNWGVDDVESFPTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRIPIQTL 381  
 QY 361 ROELSPLYPSPGTPLEVPKPGOPSEPHGSPGQDTFCQKADGLYPNPRSSSFVSCA 420  
 DB 382 ROELSPLYPSPGTPLEVPKPGOPSEPHGSPGQDTFCQKADGLYPNPRSSSFVSCA 441  
 QY 421 AGRLFQOCSPTGLVFNNSCKCCTWN 445  
 DB 442 AGRLFQOCSPTGLVFNNSCKCCTWN 466

## RESULT 2

Q9H3V8 ID Q9H3V8 PRELIMINARY; PRT; 387 AA.  
 AC Q9H3V8;  
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
 DE Chitotriosidase precursor.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96064695; PubMed=7592832;  
 RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;  
 RT "Cloning of a cDNA encoding chitotriosidase, a human chitinase  
 produced by macrophages";  
 RL J. Biol. Chem. 270:26252-26256(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98421482; PubMed=9748235;  
 RA Boot R.G., Renkema G.H., Verhoeck M., Strijland A., Blik J.,  
 RA de Meuleneester T.M., Mannens M.M., Aerts J.M.;  
 RT "The human chitotriosidase gene. Nature of inherited enzyme  
 deficiency";  
 RL J. Biol. Chem. 273:25680-25685(1998).  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 CC EMBL; U62662; AAG10644.1; -;  
 DR InterPro; IPR001579; Chitinase\_18/2.  
 DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
 DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
 DR PROSITE; PS01095; CHITINASE\_18; 1.  
 KW Glycosidase; Hydrolase; Signal.  
 FT SIGNAL  
 FT CHAIN 1 21  
 FT SEQUENCE 387 AA; 43133 MW; 03A272B8BC5E0D71 CRC64;  
 SQ

Query Match 80.7%; Score 1935; DB 4; Length 387;  
 Best Local Similarity 99.7%; Pred. No. 6.8e-142;  
 Matches 364; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHLSTTEWNETLYQEFN 60  
 DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHLSTTEWNETLYQEFN 81  
 QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEY 120  
 DB 82 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEY 141  
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAVPAQOTYVDAGYEDVKIA 180  
 DB 142 PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAVPAQOTYVDAGYEDVKIA 201  
 QY 181 QNLDFVNLMAVDPHGSWEKVTGHSNPLKYRQESGAAASLNVDAAVQOVLQKGTTPASKLI 240  
 DB 202 QNLDFVNLMAVDPHGSWEKVTGHSNPLKYRQESGAAASLNVDAAVQOVLQKGTTPASKLI 261  
 QY 241 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGMLAYEVCWSKGGATKQRIQDQK 300  
 DB 262 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGMLAYEVCWSKGGATKQRIQDQK 321  
 QY 301 VPYIFRDQNWGVDDVESFPTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRIPIQTL 360  
 DB 322 VPYIFRDQNWGVDDVESFPTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRIPIQTL 381  
 QY 361 ROELS 365  
 DB 382 ROELN 386

## RESULT 3

Q9D7Q1 ID Q9D7Q1 PRELIMINARY; PRT; 396 AA.  
 AC Q9D7Q1;  
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
 DE 2300002L19RIK protein.  
 GN 2300002L19RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=TONGUE;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner T., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann J., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 CC EMBL; AK009012; BAB26025.1; -;

```
DR MGD; MGI:1919134; 2300002L191K.
DR InterPro; IPR000677; 2S Globulin.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00551; 2SGLOBULIN.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 396 AA; 43413 MW; 32BB674B61AF7123 CRC64;

Query Match 55.6%; Score 1334; DB 11; Length 396;
Best Local Similarity 76.4%; Pred. No. 2.9e-95;
Matches 249; Conservative 26; Mismatches 49; Indels 2; Gaps 1;

QY 1 AKLVCFYTNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTHQSLSTTEWNETLYQEFN 60
DB 22 AKLVCFYTNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTHQSLSTTEWNETLYQEFN 81
QY 61 GLKXNPKLKTLLAIGGMNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPFGLDLDMWEY 120
DB 82 SLKTNPKLKTLLAIGGMNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPFGLDLDMWEY 141
QY 121 PGSQSPAVDKERFTLVODLANAFOEAEQATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 180
DB 142 PGSQSPAVDKERFTLVODLANAFOEAEQATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 201
QY 181 QNLDVFNLMAYDPHGSWEKVTGNSPLYKROESGAASLNVDAAVOQLKGTTPASKLI 240
DB 202 QNLDVFNLMAYDPHGSWEKVTGNSPLYKROESGAASLNVDAAVOQLKGTTPASKLI 261
QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPKTEGGMLAYVEVCWSW-KGATKQRIQDQ 300
DB 262 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPKTEGGMLAYVEVCWSW-KGATKQRIQDQ 319
QY 301 VPIYIFRDNQWGFDDVESPKTKVSYL 326
DB 320 VPIYIFRDNQWGFDDVESPKTKVSYL 345

RESULT 4
Q9BZP6 PRELIMINARY; PRT; 476 AA.
AC Q9BZP6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative chitinase precursor [SC 3.2.1.14].
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=STOMACH, AND LUNG;
RX MEDLINE=21125893; PubMed=11085997;
RA Boot R.G., Blommaert E.F.C., Swart E., Ghaubarali-Van Der Vlugt K.,
RA Bijl N., Moe C., Place A., Aerts J.M.P.G.;
RT Identification of a Novel Acidic Mammalian Chitinase Distinct from
RT Chitotriosidase."
RL J. Biol. Chem. 276:6770-6778 (2001).
DR EMBL; AF290004; AAG60019.1; -.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2_1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 21 POTENTIAL.
SQ SEQUENCE 476 AA; 52271 MW; 92B27BAD2F7EB4CC CRC64;

Query Match 55.2%; Score 1323; DB 4; Length 476;
Best Local Similarity 52.2%; Pred. No. 2.8e-94;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 KLVCFYTNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTHQSLSTTEWNETLYQEFN 61
DB 23 QLVCFYTNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTHQSLSTTEWNETLYQEFN 82
QY 62 LKXNPKLKTLLAIGGMNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPFGLDLDMWEY 121
DB 83 LKXNPKLKTLLAIGGMNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPFGLDLDMWEY 142
QY 122 GSQSPAVDKERFTLVODLANAFOEAEQATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 181
DB 143 GSQSPAVDKERFTLVODLANAFOEAEQATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 202
QY 182 NLDVFNLMAYDPHGSWEKVTGNSPLYKROESGAASLNVDAAVOQLKGTTPASKLI 241
DB 203 NLDVFNLMAYDPHGSWEKVTGNSPLYKROESGAASLNVDAAVOQLKGTTPASKLI 262
QY 242 GMPTYGRSFTLLASSDTRVGAPATGSGTPGPKTEGGMLAYVEVCWSW-KGATKQRIQDQ 299
DB 263 GMPTYGRSFTLLASSDTRVGAPATGSGTPGPKTEGGMLAYVEVCWSW-KGATKQRIQDQ 322
QY 300 KVPYIFRDNQWGFDDVESPKTKVSYLQKGLGAMWALDLDLPAGFCNQGRIYLIQT 359
DB 323 KVPYIFRDNQWGFDDVESPKTKVSYLQKGLGAMWALDLDLPAGFCNQGRIYLIQT 382
QY 360 LRQELSLPYLPSGTPELVPKPGQSEPP-
DB 383 LRQELSLPYLPSGTPELVPKPGQSEPP-
QY 405 GLYPNPRSSSFYSACAGRLFOOSCTGTGVNSCKCCTW 444
DB 436 GLYPNPRSSSFYSACAGRLFOOSCTGTGVNSCKCCTW 475

RESULT 5
Q9JLN1 PRELIMINARY; PRT; 472 AA.
AC Q9JLN1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative chitinase precursor (Fragment).
GN CHIA OR YNL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC1476668; TISSUE=SKIN;
RA Price P.A., Harris S.C., Williamson M.K.;
RT "YNL, A Putative Mouse Chitinase."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY; BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; AF154571; AAF31644.1; -.
DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2_1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase; Signal.
FT NON_TER 1 1
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 472 PUTATIVE CHITINASE.
```



RA Suzuki M., Morimatsu M., Yamashita T., Iwanaga T., Syuto B.;  
 RT "A novel serum chitinase that is expressed in bovine liver.";  
 RL FEBS Lett. 506:127-130(2001).  
 DR EMBL: AB051629; BAB71805.1; --  
 DR InterPro; IPR001579; Chitinase\_18/2.  
 DR InterPro; IPR002557; Chitin\_bind\_Pera.  
 DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR Pfam; PF01607; CBM\_14; 1.  
 DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
 DR PROSITE; PS01095; CHITINASE\_18; UNKNOWN\_1.  
 DR SIGNAL 1 18  
 FT CHAIN 19 488 TOAD PANCREATIC CHITINASE.  
 SQ SEQUENCE 472 AA; 52129 MW; 74A600E8DA04B1E CRC64;

Query Match 54.4%; Score 1304; DB 6; Length 472;  
 Best Local Similarity 52.2%; Pred. No. 8.2e-93;  
 Matches 239; Conservative 72; Mismatches 123; Indels 24; Gaps 5;  
 QY 2 KLVCFYFTWNAQVROGEARFLPKDLPSCILTHIYAFAGMTNHQSLSTTEWDDTLTVOEFNG 61  
 DB 23 QLVCFYFTWNAQVROGEARFLPKDLPSCILTHIYAFAGMTNHQSLSTTEWDDTLTVOEFNG 82  
 QY 62 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 121  
 DB 83 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 142  
 QY 122 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQVTVYDAGYVDKIAQ 181  
 DB 143 GFRGSPQDHLFTLVQETRAFQEAQTSKGERLLLSAAVPAQVTVYDAGYVDKIAQ 202  
 QY 182 NLDVFNLMAYDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLQKTPASKLIL 241  
 DB 203 YLDFTHVMTDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLQKTPASKLIL 262  
 QY 242 GMPTYGRSFTLASSSDTRVGPATGSGTPGPTKEGGLAYVEVCW--KGATKQRIQDO 299  
 DB 263 GPPAYGHNFILRDSNNGIGAPTSGAGPAGPYTRAGFWAYYEICAFKLDGATEAWDDSQ 322  
 QY 300 KVPYIFRDQWGFDDVESFVKVSKYKQGLGGAMVWALDLDLDFAGFCNCGRYPLIQT 359  
 DB 323 NVPYAYKGTWGVYDYNVSFRKAKWLKNNFGGAMVWALDLDLDFAGFCNCGRYPLINT 382  
 QY 360 LRQELSLVPLSGTPELEVPKQSPSEPHGSPGQDT-----FCQKADGL 406  
 DB 383 LKDALG--LKSATCNAST-----QSSEP--NSSPGNESGKNSSSSEGRGYCAGKADGL 433  
 QY 407 YNPFRSSFYSCAGRLFQOQSCPTGLVFSNCKCTW 444  
 DB 434 YPVDNRNAPWNCVNGITYKQNLCTGLVFDTSCHCCNW 471

Query Match 54.4%; Score 1304; DB 6; Length 472;  
 Best Local Similarity 52.2%; Pred. No. 8.2e-93;  
 Matches 239; Conservative 72; Mismatches 123; Indels 24; Gaps 5;

QY 2 KLVCFYFTWNAQVROGEARFLPKDLPSCILTHIYAFAGMTNHQSLSTTEWDDTLTVOEFNG 61  
 DB 23 QLVCFYFTWNAQVROGEARFLPKDLPSCILTHIYAFAGMTNHQSLSTTEWDDTLTVOEFNG 82  
 QY 62 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 121  
 DB 83 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 142  
 QY 122 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQVTVYDAGYVDKIAQ 181  
 DB 143 GFRGSPQDHLFTLVQETRAFQEAQTSKGERLLLSAAVPAQVTVYDAGYVDKIAQ 202  
 QY 182 NLDVFNLMAYDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLQKTPASKLIL 241  
 DB 203 YLDFTHVMTDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLQKTPASKLIL 262  
 QY 242 GMPTYGRSFTLASSSDTRVGPATGSGTPGPTKEGGLAYVEVCW--KGATKQRIQDO 299  
 DB 263 GPPAYGHNFILRDSNNGIGAPTSGAGPAGPYTRAGFWAYYEICAFKLDGATEAWDDSQ 322  
 QY 300 KVPYIFRDQWGFDDVESFVKVSKYKQGLGGAMVWALDLDLDFAGFCNCGRYPLIQT 359  
 DB 323 NVPYAYKGTWGVYDYNVSFRKAKWLKNNFGGAMVWALDLDLDFAGFCNCGRYPLINT 382  
 QY 360 LRQELSLVPLSGTPELEVPKQSPSEPHGSPGQDT-----FCQKADGL 406  
 DB 383 LKDALG--LKSATCNAST-----QSSEP--NSSPGNESGKNSSSSEGRGYCAGKADGL 433  
 QY 407 YNPFRSSFYSCAGRLFQOQSCPTGLVFSNCKCTW 444  
 DB 434 YPVDNRNAPWNCVNGITYKQNLCTGLVFDTSCHCCNW 471

## RESULT 8

Q90W34 PRELIMINARY; PRT; 488 AA.  
 AC Q90W34;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Toad pancreatic chitinase (EC 3.2.1.14).  
 GN TPCASE.  
 OS Bufo japonicus (Japanese toad).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufonidae; Bufo.  
 OX NCBI\_TaxID=8387;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PANCREAS;  
 RA Oshima H.;  
 RL Thesis (2001), Department of Department of Biology, Waseda University,  
 RL Tokyo, Japan.  
 DR EMBL: AJ345054; CAC87888.1; --  
 DR InterPro; IPR002086; Aldehyde\_dehydr.  
 DR InterPro; IPR001579; Chitinase\_18/2.  
 DR InterPro; IPR002557; Chitin\_bind\_Pera.

DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR Pfam; PF01607; CBM\_14; 1.  
 DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
 DR PROSITE; PS01095; CHITINASE\_18; UNKNOWN\_1.  
 DR Glycosidase; Hydrolase; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 488 TOAD PANCREATIC CHITINASE.  
 SQ SEQUENCE 488 AA; 54319 MW; E75E1AD3CB2B4919 CRC64;

Query Match 54.3%; Score 1302.5; DB 13; Length 488;  
 Best Local Similarity 51.0%; Pred. No. 1.1e-92;  
 Matches 240; Conservative 74; Mismatches 126; Indels 31; Gaps 6;

QY 2 KLVCFYFTWNAQVROGEARFLPKDLPSCILTHIYAFAGMTNHQSLSTTEWDDTLTVOEFNG 61  
 DB 20 KLVCFYFTWNAQVROGEARFLPKDLPSCILTHIYAFAGMTNHQSLSTTEWDDTLTVOEFNG 79  
 QY 62 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 121  
 DB 80 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 139  
 QY 122 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQVTVYDAGYVDKIAQ 181  
 DB 140 GSRGSPEDKQFTVLIOEMLDAPNEEARSSGLPRLLITAAVSAAGKGTIDAGYEIAKIQ 199  
 QY 182 NLDVFNLMAYDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLQKTPASKLIL 241  
 DB 200 LLDFTVMTDFHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLQKTPASKLIL 259  
 QY 242 GMPTYGRSFTLASSSDTRVGPATGSGTPGPTKEGGLAYVEVCW--KGATKQRIQDO 300  
 DB 260 GFTYGRFTRFNPNCMDVGPVSGAGSAGPYTRAGFWAYYEICATLSSGTVKWPIDOR 319  
 QY 301 VPIYIFRDQWGFDDVESFVKVSKYKQGLGGAMVWALDLDLDFAGFCNCGRYPLIQT 360  
 DB 320 VPIYACKSNEWGFDDVESFVKVSKYKQGLGGAMVWALDLDLDFAGFCNCGRYPLIQT 379  
 QY 361 RQEL-----SLPLPSGT-----PELEVPKQSPSEPHGSPGQD 396  
 DB 380 KSLLEGSTVNCPEICGGISITPTPSATTTTTTTAKPDCTTPEP--PVTTPVPV-PVID 436  
 QY 397 ---TPCQKADGLYNPNRERSFYSCAGRLFQOQSCPTGLVFSNCKCTW 444  
 DB 437 VDPNFCVEKTDGLHVNPLNTNKFYICANGRTYSMKADGLVFAQSCNCCW 487

## RESULT 9

Q99PH2 PRELIMINARY; PRT; 473 AA.  
 AC Q99PH2;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Acidic mammalian chitinase (EC 3.2.1.14).  
 GN CHIA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALE/C; TISSUE=LUNG;  
 RX MEDLINE=21125893; PubMed=11085997;  
 RA Boot R.G.; Blommaert E.F.C.; Swart E.; Ghauharali-Van Der Vlught K.,  
 RA Bijl N.; Moe C.; Place A.; Aerts J.M.F.G.;  
 RT "Identification of a Novel Acidic Mammalian Chitinase Distinct from  
 RT Chitotriosidase";  
 RL J. Biol. Chem. 276:6770-6778(2001).  
 CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 CC HYDROLASES).  
 DR EMBL: AF290003; AAG60018.1; --

```
DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 473 AA; 51977 MW; 389D87557B8C0784 CRC64;

Query Match 54.3%; Score 1301.5; DB 11; Length 473;
Best Local Similarity 51.5%; Pred. No. 1.3e-92;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LVCYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLSTTEWNDFTLYQEFNG 62
DB 1 LVCYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLSTTEWNDFTLYQEFNG 61
QY 63 KKNMPLKLTLLAIGGNWFGTKQFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 122
DB 84 KNRNPKLTLLAIGGNWFGTKQFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 121
QY 123 SQGSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 182
DB 144 SRGSPQDKHLFTVLVKEMREAEFEQAEIESNRPLAVTRAAGGIGNIQAGYIEPELSKY 203
QY 183 LDFVNLMAVDYFHGSWEKVTGCHNSPLYKROESGAASLNVDAAVQOVLKQGTTPASKLI 242
DB 204 LDFIHWMTYDLHGSWEKVTGCHNSPLYKROESGAASLNVDAAVQOVLKQGTTPASKLI 241
QY 243 MPTYGRSFTLASSSDTRVGAPATGTPGPTKEGGMALAYEVCWS--KGATKQRIODOK 300
DB 264 FPEYGHFTLRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYEICTFLRSAGTEWVNDASQ 323
QY 301 VPIYFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDPAFCSCNQGRYPLIOTL 360
DB 324 VPYAYKANLEWLGVDNIKFSVKAQWLKQNNFGGAMIWALDLDFTGSCDQKGFPLTSTL 383
QY 361 RQELSLPYLPSGTPELEVPKPGQSPSE-----EHSQSPQDQTFCCQKADGLYP 408
DB 384 NKALGISTEGCTAPDV-----PSEFVTTPGSGSGSGSGSGSGSGSCGFCADKADGLYP 436
QY 409 NPRESRFSYSCAAGRLFQOQSCPTGLVFSNSCKCTW 444
DB 437 VADDRNAPWQINGITVYQHCQAGLVFDTSNCN 472

RESULT 10
Q9WTV1 PRELIMINARY; PRT; 352 AA.
AC Q9WTV1
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycoprotein-39 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEWIS;
RA Wending U.; Boots A.M.H.; van Eden W.;
RT "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
RT potential autoantigen in arthritis."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF062038; AAD22610.1; -
DR HSSP; P07254; 1CTN.
DR InterPro; IPR000677; 2S Globulin.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.

DR PRINTS; PRO0551; 2S GLOBULIN.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR NON_TER 1
FT NON_TER 352
FT NON_TER 352
SQ SEQUENCE 352 AA; 39391 MW; CBDE991610AC936C CRC64;

Query Match 43.9%; Score 1053.5; DB 11; Length 352;
Best Local Similarity 55.5%; Pred. No. 1.4e-73;
Matches 198; Conservative 63; Mismatches 87; Indels 9; Gaps 5;

QY 2 KLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLSTTEWNDFTLYQEFNG 61
DB 2 KLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLSTTEWNDFTLYQEFNG 61
QY 62 LKKNMPLKLTLLAIGGNWFGTKQFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 121
DB 62 LKTRNPNLKTLLSVGVNFGSGRFSRISVNAKSRKTFVQSVAPFLRTYGFDCGLDLAWLP 121
QY 122 GSQSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 181
DB 122 GPK-----DKQHFTTLIKELKAEFTKEVQ--PQTEKLLLSAAVSAVSGKVTLDSCGYDVAQIA 175
QY 182 NLDVNLMAVDYFHGSWEKVTGCHNSPLYKROESGAASLNVDAAVQOVLKQGTTPASKLI 241
DB 176 HUDFNLMTYDFHGTWRHTTGHHSPLFRGQDTPDRFSNVVDYGVGYMLRLGAPTNNKLYM 235
QY 242 GMPTYGRSFTLASSSDTRVGAPATGTPGPTKEGGMALAYEVCWS--KGATKQRIODOK 300
DB 236 GIPTGKGFSTLA--SENVGAPISGGLPGRYKKEGLAYEICDFLRCGAEVHRLGQ 294
QY 301 VPIYFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDPAFCSCNQGRYPL 356
DB 295 VPFATKGNQWGVDDPESVKNKVKYLNKQLAGAMVWALDLDPRGSCFCHNVHPL 351

RESULT 11
Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to chitinase 3-like 1 (cartilage glycoprotein-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008568; AAH08568.1; -
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN 1.
SQ SEQUENCE 383 AA; 42625 MW; 6C3EF133BDC7C2D1 CRC64;

Query Match 43.6%; Score 1046.5; DB 4; Length 383;
Best Local Similarity 53.6%; Pred. No. 5.6e-73;
Matches 196; Conservative 66; Mismatches 95; Indels 9; Gaps 5;

QY 2 KLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLSTTEWNDFTLYQEFNG 61
DB 23 KLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLSTTEWNDFTLYQEFNG 82
QY 62 LKKNMPLKLTLLAIGGNWFGTKQFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 121
DB 83 LKRNPNLKTLLSVGVNFGSGRFSKIASNTQSRFTFKSVPPFLRTHGFCGLDLAWLP 142
QY 122 GSQSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 181
```

```
Db 143 GRR-----DKOHFTTLIKEMKAFTKEAQ-PGKKQLLLSAAALSAGKVTIDSSYDIATKISQ 196
QY 182 NUDFVNLMAFYDFHGSWEKVTGNSPLKYQESGAAASLNVDAAVOOVLQKTPASKLIL 241
Db 197 HLDFTSLTYDFHGAWRGTTHHSPFLFRQEDASDRFSNTDYAVGYMLRLGAPASKLVM 256
QY 242 GMPTYGRSFTLASSSDTRVGAPATSGTGPPTKEGGMLAYVEVCSW-KGATKQRIQDQK 300
Db 257 GIPTGRSFTLA-SSETGVGAPISGPIGRFTKEAGTLAYYEICDFLRGATVHRILGQQ 315
QY 301 VPIYFRDNQWGFDDVESFKTKVSKYKQGLGGMVWALDLDLDDPAGFSCNQG-RYPLIQ 359
Db 316 VPYATKGNQWGYDDQESVKSVQVYLKQRLAGAMVWALDLDLDDPAGFSCNQGLRPLTNA 375
QY 360 LRQELS 365
Db 376 IXDALA 361

RESULT 12
Q96F97 PRELIMINARY; PRT; 390 AA.
AC Q96F97;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Chitinase 3-like 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011460; AAH11460.1; -.
DR InterPro; IPR001223; Glyco_hydro.18.
DR Pfam; PF00704; Glyco_hydro.18; 1.
DR ProDom; PD000471; Glyco_hydro.18; 1.
SQ SEQUENCE 390 AA; 43529 MW; 6801E4AF721AD63F CRC64;

Query Match 43.4%; Score 1041.5; DB 4; Length 390;
Best Local Similarity 51.9%; Pred. No. 1.4e-72;
Matches 190; Conservative 76; Mismatches 91; Indels 9; Gaps 4;

QY 2 KLVCYFTNWAQYRQGEARFLPKDLPSCLTLLIYAFAGTNNHQLSTTEWNETLYQEFNG 61
Db 28 KLVCYFTNWSQDRQPGKFTPENIDPFLCSHLIYGFASITENKVIKDKSEVMLYQTINS 87
QY 62 LKKNPKLKTLLAIGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDCGLDLDWEYP 121
Db 88 LKTKNPKLKLISLIGLYLFGFHPMDSSSTRLEFINSILFLRNHNFGLDVSWIYP 147
QY 122 GSQGSFPAVDKERFTTLVODLANAFOEAQTSKGERLLLSAAVPAQOTYVDAGYVDKIAQ 181
Db 148 DQK-----ENTHTVLIHLAEAFQKFTKSTKERLLLTIVGSAGROMIDNSYQVEKLAK 202
QY 182 NUDFVNLMAFYDFHGSWEK--VTGNSPLKYQESGAAASLNVDAAVOOVLQKTPASKL 239
Db 203 DLDFTNLLSDFHGSWEKPLITGHNSPLSKQDRGPSSYYNVEVAVGYWIHKGPSEKV 262
QY 240 ILGMPTYGRSFTLASSSDTRVGAPATSGTGPPTKEGGMLAYVEVCSW-KGATKQRIQD 298
Db 263 VMGIPTYGRSFTLA-SAETTVGAPASGPGAAGPITSSGFLAYYEICQFLKAKITRLQD 321
QY 299 QKVPYIFRDNQWGFDDVESFKTKVSKYKQGLGGMVWALDLDLDDPAGFSCNQGRYPLIQ 358
Db 322 QQVPAVKNQWGVDDVKSMEIKVQFLKNLNLGGMVWISIDMDDDFTTGKSCNQGPYPLVQ 381
QY 359 TURQEL 364
::: |
```

```
Db 382 AVGRSL 387

RESULT 13
Q8SPQO PRELIMINARY; PRT; 383 AA.
AC Q8SPQO;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE BP40 precursor.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY GLAND;
RA Paramasivam M., Saravanan K., Sharma S., Mohanty A.K., Yadav S.,
RA Srinivasan A., Singh T.P.;
RT "Goat mammary gland mRNA for a novel protein, BP40; complete coding
region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY081150; AAL87007.1; -.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 383 AA; 42893 MW; 17655ED4BE4B9F5F CRC64;

Query Match 42.9%; Score 1029.5; DB 6; Length 383;
Best Local Similarity 51.9%; Pred. No. 1.2e-71;
Matches 190; Conservative 69; Mismatches 98; Indels 9; Gaps 5;

QY 2 KLVCYFTNWAQYRQGEARFLPKDLPSCLTLLIYAFAGTNNHQLSTTEWNETLYQEFNG 61
Db 23 KLCYVTSWSQYREGDGCSPDAIDPFLCTTHIYSPANISNEIDTWEMNDVTLTLNT 82
QY 62 LKKNPKLKTLLAIGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDCGLDLDWEYP 121
Db 83 LKNRPKLTLLSVGWNFGPGRFSKIAKTSQSRFTFKSVPPFLRTHGFGDLDLAWLYP 142
QY 122 GSQGSFPAVDKERFTTLVODLANAFOEAQTSKGERLLLSAAVPAQOTYVDAGYVDKIAQ 181
Db 143 GRR-----DKRHLTGLVKEMKAEPAREAQ-AGTERLLLSAAVSAKIAIDRCYDIAQISR 196
QY 182 NUDFVNLMAFYDFHGSWEKVTGNSPLKYQESGAAASLNVDAAVOOVLQKTPASKLIL 241
Db 197 HLDFTSLTYDFHGAWRQTVGHSPFLFRQEDASDRFSNADYAVSYMLRLGAPANKLVM 256
QY 242 GMPTYGRSFTLASSSDTRVGAPATSGTGPPTKEGGMLAYVEVCSW-KGATKQRIQDQK 300
Db 257 GIPTGRSFTLASSK-TDVGAPISGPIGRFTKEGILAYYEICDFLHGATTHFRDQO 315
QY 301 VPIYFRDNQWGFDDVESFKTKVSKYKQGLGGMVWALDLDLDDPAGFSCNQG-RYPLIQ 359
Db 316 VPYATKGNQWAVDDQESVKNKARYLKNRQLAGAMVWALDLDLDDPGRGTFGQNLTFPLTSA 375
QY 360 LRQELS 365
Db 376 VKDVLA 381
::: |

RESULT 14
Q29411 PRELIMINARY; PRT; 383 AA.
AC Q29411;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE 38 kDa heparin-binding glycoprotein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
```

NCBI\_TaxID=9823;  
 [1]  
 RX SEQUENCE FROM N.A.  
 RP TISSUE=SMOOTH MUSCLE;  
 RC MEDLINE=95286589; PubMed=7768902;  
 RA Shackleton L.M., Mann D.M., Millis A.J.;  
 RT "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in  
 RT differentiating vascular smooth muscle cells as a member of a group of  
 RT proteins associated with tissue remodeling.";  
 RL J. Biol. Chem. 270:13076-13083(1995).  
 RN [2]  
 RX SEQUENCE FROM N.A.  
 RP TISSUE=SMOOTH MUSCLE;  
 RC Shackleton L.M., Mann D.M., Millis A.J.T.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U19900; AAA86482.1; -;  
 DR EMBL: Z47803; CAA8764.1; -;  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 DR ProDom: PD000471; Glyco\_hydro\_18; 1.  
 KW Heparin-binding.  
 SQ SEQUENCE 383 AA; 42443 MW; 3D1039F49910BDC4 CRC64;  
 Query Match 42.6%; Score 1021.5; DB 6; Length 383;  
 Best Local Similarity 52.2%; Pred. No. 4.9e-71;  
 Matches 191; Conservative 67; Mismatches 99; Indels 9; Gaps 5;  
 RX 2 KLVCYFTNNAOYRQGEARFLPKDLPSCLTHTLIYAFAGTNHOLSTENDETLYQFNG 61  
 DB 23 KLVCYFTNNAOYRQGEARFLPKDLPSCLTHTLIYAFAGTNHOLSTENDETLYQFNG 82  
 RX 62 LKQNPENKLTLLAIGGNFQGTQFTDMVATANNROTFFVNSAIRFLRKYSPFGDGLDWEYP 121  
 DB 83 LKQNPENKLTLLSAGVGNFQSGFRKSIASNTQSRFTFKSVPPFLTHGDFGLDLAWISP 142  
 RX 122 GSGGSPAVDKERTFTLLVQDLANAFQEAQTSKGERLLLSAANPAGQTYVDAGVEVDKIAQ 181  
 DB 143 GR- - - -DKRHLLTTLVKEMKAEFVREA-LPGTTERLLSGAVSAGKVAIDRGYDIAQISQ 196  
 RX 182 NLDFVNLMAYDFHSGSEKVTGHSNPLYKQESGAASLNVDAAVQVQKQKGPASKLLIL 241  
 DB 197 HLDFTISLLTYDFHGAQRQTTHSHPLFRGGQDASSRDNFNAVSVYLRGLGAPANKLVN 256  
 RX 242 GMPYIGRSFTLASSSDTRVGAPATSGTGPFTKEGMLAYEVCWSW-KGATKQRIQDQK 300  
 DB 257 GIFTGSRFTLASSK-TDVGAPASGFGIPGRFTKEGILAYEICDFLQGATVRRPLGQQ 315  
 RX 301 VPYIFRDQNVGPDVSEFVKTVSYLKQKGLGAMVWALDLDLDFAGFSCNQG-RYPLIQT 359  
 DB 316 VPYATKGNQWVGDDQESVKNKAKYLSKRLAGAMVWTLDDLDFRGNFCQNLRPLTSA 375  
 RX 360 LRQELS 365  
 DB 376 IKDVL 381  
 RESULT 15  
 Q9VZV2 PRELIMINARY; PRT; 696 AA.  
 AC Q9VZV2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE CG1869 protein.  
 GN CG1869.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;

MEDLINE=20196006; PubMed=10731132;  
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Zheng L.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RL "The genome sequence of Drosophila melanogaster.";  
 RT Science 287:2185-2195(2000).  
 RX EMBL: AE003477; AAF47714.1; -;  
 DR FlyBase; FBgn035398; CG1869.  
 DR InterPro: IPR001579; Chitinase\_18/2.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 2.  
 DR ProDom: PD000471; Glyco\_hydro\_18; 2.  
 DR PROSITE: PS01095; CHITINASE\_18; UNKNOWN 1.  
 SQ SEQUENCE 696 AA; 77366 MW; 6877585837A19CF6 CRC64;  
 Query Match 42.3%; Score 1015.5; DB 5; Length 696;  
 Best Local Similarity 53.0%; Pred. No. 3.6e-70;  
 Matches 197; Conservative 52; Mismatches 110; Indels 13; Gaps 5;  
 RX 2 KLVCYFTNNAOYRQGEARFLPKDLPSCLTHTLIYAFAGTNHOLSTENDET- - - -L 55  
 DB 124 KIVCYTNNQSVYRKGIVFVPEIPADLCTHIFAFGLKKNKLSYESNDETKNVPLG 183  
 RX 56 YQEFNGLKKNPKLTLAIGGNFQGTQFTDMVATANNROTFFVNSAIRFLRKYSPFGDGLD 115  
 DB 184 YERMVTLKANPKLLALGCGSFGTKQFKDMSSTRYTRQTFVYSAIPFLKRGDFGLD 243  
 RX 116 LDWEYFGSGSPAVDKERTFTLLVQDLANAFQEAQTSKGERLLLSAANPAGQTYVDAGVE 175  
 DB 244 MDWEYF- -KGSD- -DKQNFVLLKELREAFEAQELKKPRLLLSAANPAGVDPNIRGGYD 299  
 RX 176 VDKIAQNLDVNLMAVDHSGSEKVTGHSNPLYKQESGAASLNVDAAVQVQKQKGP 235  
 DB 300 VPAIASYLDLFINLMAYDFHGRWERTGHNAPLYASTSEWRKQSLVSNAAALWVMGAP 359  
 RX 236 ASKLILGMPYIGRSFTLASSSDTRVGAPATSGTGPFTKEGMLAYEVCWSW- -KGATK 293  
 DB 360 KEKLIVGMPYIGRSFTLANPDKHGPNAPASGGREGVYTKEGFLAYEICEMLNGAVY 419  
 RX 294 QRIQDQKVPYIFRDQNVGPDVSEFVKTVSYLKQKGLGAMVWALDLDLDFAGFSCNQ- 352



Db 420 VWDDEMKVFLVDGQWVGFDDEAIRNKKHWIKSNGFGAMVWTIDMDDFKGEVCGNV 479  
Qy 353 RYPLJOTLRQEL 364  
Db 480 KYPLIGAMREEL 491

Search completed: June 29, 2003, 21:01:47  
Job time : 26.0148 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:50:19 ; Search time 30.2885 Seconds  
(without alignments)  
2001.720 Million cell updates/sec

Title: US-10-004-219B-14

Perfect score: 2506

Sequence: 1 YQLTCYFTNWAQYRPLGRF.....QQNCOAGLVFTSDCCNWA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 3

Maximum DB seq length: 3000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A. Geneseq 101002:\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2506	100.0	455	22	AAAG62541 Disease treatment
2	2506	100.0	476	22	AAAG62543 Disease treatment
3	2018	80.5	368	23	ABBY7580 Novel human protei
4	1415	56.5	377	22	AAAG62545 Disease treatment
5	1415	56.5	398	22	AAAG62544 Disease treatment
6	1323	52.8	466	18	AAW08584 Human 50 kDa chiti
7	1323	52.8	466	19	AAW04259 Human chitinase pr
8	1323	52.8	466	20	AAW42425 MO-218 clone of hu
9	1323	52.8	466	22	AAW04322 Human chitinase pr
10	1323	52.8	466	23	ABBY76291 Human chitinase.

11	1317	52.6	466	19	AAW40260 Human chitinase pr
12	1317	52.6	466	20	AAW42426 MO-13B clone of hu
13	1317	52.6	466	22	AAE00433 Human chitinase pr
14	1317	52.6	466	23	ABBY76292 Human chitinase.
15	1316	52.5	466	18	AAW31498 Human chitotriosid
16	1297	51.8	464	18	AAW31497 Human chitotriosid
17	1177	47.0	373	19	AAW40262 Human chitinase pr
18	1177	47.0	373	19	AAW40261 Human chitinase pr
19	1177	47.0	373	20	AAW42427 Clone of the C-ter
20	1177	47.0	373	20	AAW42428 Chitinase amino ac
21	1177	47.0	373	23	ABBY76293 Human chitinase tr
22	1177	47.0	373	23	ABBY76294 Human chitinase an
23	1173	46.8	387	18	AAW08585 Human 39 kDa chiti
24	1011.5	40.4	718	16	AAW73992 Murine oviduct spe
25	1008.5	40.2	696	22	ABBS8615 Drosophila melanog
26	995	39.7	203	22	AAW92862 Human digestive ey
27	978.5	39.0	421	19	AAW53122 Amino acid sequenc
28	970.5	38.7	385	19	AAW53123 Amino acid sequenc
29	970.5	38.7	385	19	AAW47033 Human cartilage gp
30	970.5	38.7	416	19	AAW47034 Human cartilage gp
31	970.5	38.7	421	19	AAW53121 Amino acid sequenc
32	970.5	38.7	423	18	AAW35930 Human cartilage gl
33	963.5	38.4	383	18	AAW26751 Bovine whey protei
34	961.5	38.4	383	21	AAW03442 Gp8k protein sequ
35	961.5	38.4	668	16	AAW73993 Hamster oviduct sp
36	943.5	37.6	537	16	AAW73991 Bovine oviduct spe
37	890.5	35.5	4498	22	ABBS8595 Drosophila melanog
38	863	34.4	554	18	AAW01824 Manduca sexta larv
39	863	34.4	554	21	AAW07183 Manduca sexta gut
40	853	34.0	460	22	ABBS4366 Drosophila melanog
41	768.5	30.7	595	22	ABBY7137 Polypeptide involv
42	760	30.3	321	9	AAW81342 D. pteronyssius 98
43	711.5	28.4	490	21	AAW52535 D. pteronyssius 98
44	711.5	28.4	490	23	AAW96339 Der. HMW-map poly
45	711.5	28.4	509	21	AAW52533 D. pteronyssius 98

ALIGNMENTS

RESULT 1

AAAG62541

ID AAAG62541 standard; protein; 455 AA.

XX

AC AAAG62541;

XX

DT 24-AUG-2001 (first entry)

XX

DE Disease treatment related protein SEQ ID NO: 1.

XX

KW Disease treatment; infection; chronic occlusive pulmonary disease;

KW bronchial asthma.

XX

OS Homo sapiens.

XX

PN WO200136633-A1.

XX

PD 25-MAY-2001.

XX

PF 14-NOV-2000; 2000WO-JP08015.

XX

PR 15-NOV-1999; 99JP-0324467.

XX

PA (TAKE ) TAKEDA CHEM IND LTD.

XX

PI Nakanishi A, Morita S;

XX

DR WPI; 2001-397791/42.

XX

PT New proteins, peptides and DNA for treatment of bronchial asthma,

PT chronic occlusive lung disease and infectious disease

XX

PS Claim 1; Page 97-99; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be  
CC used in the treatment and prevention of infectious diseases. Inhibitors  
CC of the protein can be used to treat bronchial asthma and chronic  
CC occlusive pulmonary disease. The present sequence is a protein described  
CC in the exemplification of the invention.  
XX  
SQ Sequence 455 AA;  
Query Match 100.0%; Score 2506; DB 22; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.3e-219;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 YQLTCYFTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRONNEITTIENWDTLYQAFN 60  
DB 1 YQLTCYFTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRONNEITTIENWDTLYQAFN 60  
QY 61 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFDFGLDFDWEY 120  
DB 61 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFDFGLDFDWEY 120  
QY 121 PGRSGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSYEIPQLS 180  
DB 121 PGRSGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSYEIPQLS 180  
QY 181 QYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGNSNAYLVNDYVMYWKDNGAPAEKLI 240  
DB 181 QYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGNSNAYLVNDYVMYWKDNGAPAEKLI 240  
QY 241 VGPFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300  
DB 241 VGPFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300  
QY 301 QEPVYAYQGNVWVGYNIKSFDIKAQWLKHNKFGGAMVWAIDLDDTGTFCNQKPEPLIS 360  
DB 301 QEPVYAYQGNVWVGYNIKSFDIKAQWLKHNKFGGAMVWAIDLDDTGTFCNQKPEPLIS 360  
QY 361 TLKALGLQASCTAPAQPIETAPSPSGNGSGSSGSGSGSGSGSGSGSGSGSGSGSGSG 420  
DB 361 TLKALGLQASCTAPAQPIETAPSPSGNGSGSSGSGSGSGSGSGSGSGSGSGSGSGSG 420  
QY 421 NNRNAPFWHCVNGVTYQONCOAGLVFTSDCCNWA 455  
DB 421 NNRNAPFWHCVNGVTYQONCOAGLVFTSDCCNWA 455  
RESULT 2  
AAG62543  
ID AAG62543 standard; protein; 476 AA.  
XX  
AC AAG62543;  
DT 24-AUG-2001 (first entry)  
XX  
DE Disease treatment related protein SEQ ID NO: 5.  
XX  
KW Disease treatment; infection; chronic occlusive pulmonary disease;  
KW bronchial asthma.  
XX  
OS Homo sapiens.  
XX  
PN WO200136633-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 14-NOV-2000; 2000WO-JP08015.  
XX  
PR 15-NOV-1999; 99JP-0324467.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Nakanishi A, Morita S;  
XX

DR WPI; 2001-397791/42.  
XX  
PT New proteins, peptides and DNA for treatment of bronchial asthma,  
PT chronic occlusive lung disease and infectious disease  
XX  
PS Example 5; Page 101-103; 114pp; Japanese.  
XX  
CC The present invention provides the sequence of a protein which can be  
CC used in the treatment and prevention of infectious diseases. Inhibitors  
CC of the protein can be used to treat bronchial asthma and chronic  
CC occlusive pulmonary disease. The present sequence is a protein described  
CC in the exemplification of the invention.  
XX  
SQ Sequence 476 AA;  
Query Match 100.0%; Score 2506; DB 22; Length 476;  
Best Local Similarity 100.0%; Pred. No. 1.4e-219;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 YQLTCYFTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRONNEITTIENWDTLYQAFN 60  
DB 22 YQLTCYFTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRONNEITTIENWDTLYQAFN 81  
QY 61 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFDFGLDFDWEY 120  
DB 82 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFDFGLDFDWEY 141  
QY 121 PGRSGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSYEIPQLS 180  
DB 142 PGRSGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSYEIPQLS 201  
QY 181 QYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGNSNAYLVNDYVMYWKDNGAPAEKLI 240  
DB 202 QYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGNSNAYLVNDYVMYWKDNGAPAEKLI 261  
QY 241 VGPFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300  
DB 262 VGPFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 321  
QY 301 QEPVYAYQGNVWVGYNIKSFDIKAQWLKHNKFGGAMVWAIDLDDTGTFCNQKPEPLIS 360  
DB 322 QEPVYAYQGNVWVGYNIKSFDIKAQWLKHNKFGGAMVWAIDLDDTGTFCNQKPEPLIS 381  
QY 361 TLKALGLQASCTAPAQPIETAPSPSGNGSGSSGSGSGSGSGSGSGSGSGSGSGSGSG 420  
DB 382 TLKALGLQASCTAPAQPIETAPSPSGNGSGSSGSGSGSGSGSGSGSGSGSGSGSGSG 441  
QY 421 NNRNAPFWHCVNGVTYQONCOAGLVFTSDCCNWA 455  
DB 442 NNRNAPFWHCVNGVTYQONCOAGLVFTSDCCNWA 476  
RESULT 3  
ABB97580  
ID ABB97580 standard; protein; 368 AA.  
XX  
AC ABB97580;  
XX  
DT 27-JUN-2002 (first entry)  
XX  
DE Novel human protein SEQ ID NO: 848.  
XX  
KW Human; antianemic; vulnerary; antiinflammatory; immunomodulator;  
KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
KW expressed sequence tag.  
XX  
OS Homo sapiens.  
XX  
PN WO200222660-A2.  
XX  
PD 21-MAR-2002.  
XX

PF 10-SEP-2001; 2001WO-US26015.  
XX  
PR 11-SEP-2000; 2000US-0659671.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX  
DR WPI; 2002-292408/33.  
DR N-PSDB; ABN32766.  
XX  
An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis -  
XX  
Claim 20; SEQ ID NO 848; 509pp; English.  
XX  
The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a protein of the invention.  
XX  
SQ Sequence 368 AA;  
Query Match 80.5%; Score 2018; DB 23; Length 368;  
Best Local Similarity 100.0%; Pred. No. 3e-175;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 88 MVSTPENRTFTSVIKFLRQYEFDFGLDPDWEYPSGRGSPQDKHLFTVLVQEMREAFQ 147  
Db 1 MVSTPENRTFTSVIKFLRQYEFDFGLDPDWEYPSGRGSPQDKHLFTVLVQEMREAFQ 60  
QY 148 EAKQINKPLMTAAVAAGISIQSGYEIPQLSOYLDYIHVMTYDLHGSWEGYTGNSPL 207  
Db 61 EAKQINKPLMTAAVAAGISIQSGYEIPQLSOYLDYIHVMTYDLHGSWEGYTGNSPL 120  
QY 208 YKYPTDTSNAYLVNDVYVNYWKDNGAPAEKLI VGPPTYGHNFI LSNPNTGIGAPTSGA 267  
Db 121 YKYPTDTSNAYLVNDVYVNYWKDNGAPAEKLI VGPPTYGHNFI LSNPNTGIGAPTSGA 180  
QY 268 GPAGPYAKESGIWAYYEICTFLKNGATQGWDAPOEVVPYAYQGNVWVGYDNIKSFDIKAQW 327  
Db 181 GPAGPYAKESGIWAYYEICTFLKNGATQGWDAPOEVVPYAYQGNVWVGYDNIKSFDIKAQW 240  
QY 328 LKHNFEGGAMVWALDLDFTGTCNKGKPLISTLKALGLQSASCTAPAOPIETITRAP 387  
Db 241 LKHNFEGGAMVWALDLDFTGTCNKGKPLISTLKALGLQSASCTAPAOPIETITRAP 300  
QY 388 SSGSGSGSGSGSGSGSGGFCFCAVRANGLYPVANNRNFVHCVNGVYTYQNCQAGLVFTT 447  
Db 301 SSGSGSGSGSGSGSGSGGFCFCAVRANGLYPVANNRNFVHCVNGVYTYQNCQAGLVFTT 360  
QY 448 SCDCCNWA 455  
Db 361 SCDCCNWA 368  
RESULT 4  
AAG62545  
ID AAG62545 standard; protein; 377 AA.  
XX  
AC AAG62545;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Disease treatment related protein SEQ ID NO: 18.  
XX

Disease treatment; infection; chronic occlusive pulmonary disease;  
KW bronchial asthma.  
XX  
OS Mus sp.  
XX  
PN WO200136633-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 14-NOV-2000; 2000WO-JP08015.  
XX  
PR 15-NOV-1999; 99JP-0324467.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Nakanishi A, Morita S;  
XX  
DR WPI; 2001-397791/42.  
XX  
New proteins, peptides and DNA for treatment of bronchial asthma,  
PT chronic occlusive lung disease and infectious disease -  
XX  
PS Claim 18; Page 110-112; 114pp; Japanese.  
XX  
The present invention provides the sequence of a protein which can be  
CC used in the treatment and prevention of infectious diseases. Inhibitors  
CC of the protein can be used to treat bronchial asthma and chronic  
CC occlusive pulmonary disease. The present sequence is a protein described  
CC in the exemplification of the invention.  
XX  
SQ Sequence 377 AA;  
Query Match 56.5%; Score 1415; DB 22; Length 377;  
Best Local Similarity 68.1%; Pred. No. 3e-120;  
Matches 256; Conservative 48; Mismatches 72; Indels 0; Gaps 0;  
QY 1 YQLTCYFNWQAQYRFLGRFMPDNIDPCLCTHLIYAFAGRNNEITTIEMNDVTLYQAFN 60  
Db 1 YQLTCYFNWQAQYRFLGRFMPDNIDPCLCTHLIYAFAGRNNEITTIEMNDVTLYQAFN 60  
QY 61 GLKKNKSQLKLLAIGGNFTGAPTAMVSTPENRTFTSVIKFLRQYEFDFGLDPDWEY 120  
Db 61 GLKKNKSQLKLLAIGGNFTGAPTAMVSTPENRTFTSVIKFLRQYEFDFGLDPDWEY 120  
QY 121 PGSRSGPQDDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAGISIQSGYEIPQLS 180  
Db 121 PGSRSGPQDDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAGISIQSGYEIPQLS 180  
QY 181 QYLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTSNAYLVNDVYVNYWKDNGAPAEKLI 240  
Db 181 QYLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTSNAYLVNDVYVNYWKDNGAPAEKLI 240  
QY 241 VGFPTYGHNFI LSNPNTGIGAPTSGAGAPYAKESGIWAYYEICTFLKNGATQGWDA 300  
Db 241 VGFPTYGHNFI LSNPNTGIGAPTSGAGAPYAKESGIWAYYEICTFLKNGATQGWDA 300  
QY 301 QSVVPYAYQGNVWVGYDNIKSFDIKAQWLKHNKFGAMVWALDLDFTGTCNKGKPLIS 360  
Db 301 QSVVPYAYQGNVWVGYDNIKSFDIKAQWLKHNKFGAMVWALDLDFTGTCNKGKPLIS 360  
QY 361 TLKKGALGLQSASCTAP 376  
Db 361 TLKKGALGLQSASCTAP 376  
RESULT 5  
AAG62544  
ID AAG62544 standard; protein; 398 AA.  
XX  
AC AAG62544;  
XX  
DT 24-AUG-2001 (first entry)  
XX

DE Disease treatment related protein SEQ ID NO: 17.  
KW Disease treatment; infection; chronic occlusive pulmonary disease;  
KW bronchial asthma.

OS Mus sp.  
PN WO200136633-A1.  
PD 25-MAY-2001.  
XX 14-NOV-2000; 2000WO-JP08015.  
XX 15-NOV-1999; 99JP-0324467.  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA Nakanishi A, Morita S;  
PI WPI; 2001-397791/42.  
DR New proteins, peptides and DNA for treatment of bronchial asthma,  
PT chronic occlusive lung disease and infectious disease -  
XX Disclosure; Page 108-110; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be  
CC used in the treatment and prevention of infectious diseases. Inhibitors  
CC of the protein can be used to treat bronchial asthma and chronic  
CC occlusive pulmonary disease. The present sequence is a protein described  
CC in the exemplification of the invention.

XX Sequence 398 AA;  
Query Match 56.5%; Score 1415; DB 22; Length 398;  
Best Local Similarity 68.1%; Pred. No. 3.3e-120;  
Matches 256; Conservative 48; Mismatches 72; Indels 0; Gaps 0;

QY 1 YQLTCVFTNWAQVRPGIGRPMNDIPCLCTHLIYAFAGQNNETITWNDVTLQAFN 60  
DB 22 YQLWCYITSWAKDRPIEGSPKGNIDPCLCTHLIYAFAGQNNETITWNDVTLQAFN 81  
QY 61 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLROYEFDGLDFWYEP 120  
DB 82 GLKDKNTELKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLROYEFDGLDFWYEP 141  
QY 121 PGSRGPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAAGISNIQSGYEIPQLS 180  
DB 142 PGSRGPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAAGISNIQSGYEIPQLS 201  
QY 181 QYLDYTHVMTYDLHGSWEGYTGNSPLYKYPTDTGSAVNLVNDYVMNYKDNCAPEKLI 240  
DB 202 QSLDYIQVMTYDLHDPKDGVTGNSPLYKSPYDIGKSADLNVDVSIISYWKDHGAASEKLI 261  
QY 241 VGPFTYCHNFIILNPNSTGICAPTSGAGPAGPAGKESGIWAYVEICTFLKNGATQGWADP 300  
DB 262 VGPFAVHTFILSDPSKTGIGATITSTGPPGKYTDESGLLAYVEVCTFLNEGATEVNDAP 321  
QY 301 QEVPIYQGNVWGYDNIKSFDIKAOWLKHNFKFGAMWAIDLDDTGTFCNQGKPLIS 360  
DB 322 QEVPIYQGNVWGYDNIKSFDIKAOWLKHNFKFGAMWAIDLDDTGTFCNQGKPLIS 381  
QY 361 TLAKAGLQOSASCTAP 376  
DB 382 TLKAGDLNHSASCKGP 397

RESULT 6  
ID AAW08584  
XX AAW08584 standard; Protein; 466 AA.  
AC AAW08584;  
XX

DT 24-MAR-1997 (first entry)  
XX Human 50 kDa chitinase.  
XX Chitinase; chitotriosidase; chitin; infectious disease;  
KW gene therapy; vaccine; diagnosis; lysosomal lipodosis;  
KW Gaucher disease; leishmaniasis; sarcoidosis; multiple sclerosis;  
KW X-linked adrenoleukodystrophy; drug delivery; cosmetics; food.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH Peptide 1..21  
FT /label= Sig\_peptide  
FT Domain 178..198  
FT /label= Catalytic\_domain  
XX WO9640940-A2.  
PN 19-DEC-1996.  
XX 06-JUN-1996; 96WO-NL00225.  
XX 07-JUN-1995; 95US-0486839.  
XX (UNAM ) UNIV AMSTERDAM.  
XX Aerts JNFG;  
XX WPI; 1997-118698/11.  
DR N-PSDB; AAT50833.  
XX New human chitinase - used to treat or prevent infection by  
PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,  
PT foods, implants etc.  
XX Claim 1; Page 40-42; 58pp; English.  
XX Human chitinases of 50 kDa (AAW08584) and 39 kDa (AAW08585) are stable  
CC to many proteases, active at pH 3-8 and up to 50 deg, and stable in  
CC the circulation. They are the result of alternative splicing of  
CC RNA, and can be produced on a large-scale in transformed host cells  
CC using cDNA clones (see also AAT50833-34) obt'd. from a human macrophage  
CC library. The chitinases, optionally expressed from a gene therapy  
CC vector, are used to treat or prevent infection by chitin-contg.  
CC pathogens (e.g. fungi, protozoa, helminths) and more generally to  
CC degrade chitin. They are also useful in controlled-release drug  
CC delivery, in cosmetics, foods and dental products, for antibody  
CC prodn. and for diagnosis of diseases associated with elevated  
CC chitinase levels.

XX Sequence 466 AA;  
Query Match 52.8%; Score 1323; DB 18; Length 466;  
Best Local Similarity 52.2%; Pred. No. 1e-111;  
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;  
QY 2 QLTCTVFTNWAQVRPGIGRPMNDIPCLCTHLIYAFAGQNNETITWNDVTLQAFN 61  
DB 23 KLVCCYFTNWAQVRQGEARFLPKDLPCLCTHLIYAFAGQNNETITWNDVTLQAFN 82  
QY 62 LKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLROYEFDGLDFWYEP 121  
DB 83 LKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLROYEFDGLDFWYEP 142  
QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAAGISNIQSGYEIPQLS 181  
DB 143 GSRGSPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAAGISNIQSGYEIPQLS 202  
QY 182 YLDYTHVMTYDLHGSWEGYTGNSPLYKYPTDTGSAVNLVNDYVMNYKDNCAPEKLI 241  
DB 203 NLDYFVNLMAVDFHGSWEKVTGNSPLYKRYQESGAASLNVDAAVQOVLQKGTAPSKLIL 262

```
CC   and thus of their side effects.
XX
SQ   sequence      466 AA;

Query Match          52.8%; Score 1323; DB 19; Length 466;
Best Local Similarity 52.2%; Pred. No. 1e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY    2 QLTCTFTNWAQYPGLGRFMPDNIDPLCLTHLIYAFAGRONNBITTENNDVTLTQAENG 61
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::
Db     23 KLVCYFTFNWAQYRQEARGFLPKDLDPSLCTHLIYAFAGTNHQLSTTENDETLYQEFNG 82

QY    62 LKNKNSQLKTLLAIGWNFGCTAPFTAMWSTPENRQTFTITSVIKFLROYEPDGLDFWEYFP 121
       ||::||::||::||::||::||::||::||::||::||::||::||::||::
Db     83 LKQNNPKLKTLLAIGWNFGTQRFOTDMVAFTANNROTFFVNSAIRFLARYSPDGLDLDWEYFP 142
```

[illegible]

PR 12-MAR-1998; 98US-0039198.  
XX (ICOS-) ICOS CORP.  
XX Gray PW, Tjoelker LW;  
PI WPI; 1999-551417/46.  
DR N-PSDB; AA221847.  
XX Novel chitin-binding fragments of human chitinase used to treat fungal  
PT infections in animals  
XX Example 1; Page 58-59; 83pp; English.  
XX This is the amino acid sequence of an allelic form of the human  
CC chitinase enzyme, which is capable of degrading Chitin (a linear  
CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).  
CC The sequence also includes a Chitin binding peptide region, which lacks  
CC any chitinase activity, but which can be used to target anti-fungal  
CC agents to chitin containing fungal cells.  
CC Chitinase fragments can be used to screen for proteins or other  
CC molecules that specifically bind to the chitin-binding domain of human  
CC chitinase or that modulate its activity. These compounds are useful for  
CC immunization, as well as for purifying chitinase, as well as for  
CC detection and quantification of chitinase. Polynucleotide fragments of  
CC the invention are useful as a source of probes and primers, and to  
CC express the proteins recombinantly. The chitinase fragments, when  
CC conjugated to antifungal compounds, are used to treat animals,  
CC especially humans, infected with chitin-containing parasites such as  
CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
CC sporotrichosis, and dermatophytoses.  
CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
CC protein for treating infections, especially fungal infections, is  
CC problematic. In view of the increasing incidents of life-threatening  
CC fungal infection in e.g. immunocompromised individuals, there exists a  
CC need for identifying new compounds for treating fungal infection. The  
CC chitin-binding fragments of the present invention provide this need.  
XX Sequence 466 AA;  
Query Match 52.8%; Score 1323; DB 20; Length 466;  
Best Local Similarity 52.2%; Pred. No. 1e-111;  
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;  
2 QLTCTYTNWAQYRPGLRGMPNDIDCLCTHLYAFAGRNNEITIEWNDVLYQAFNG 61  
23 KLVCFYTNWAQYRQGEARFLPKDLPDLCTHLYAFAGMTNHLSTEWNDVLYQEFNG 82  
62 LKNNKSQLTLLAIGWNFGTAPFTAMVSTPENRQTFITSVIKFLQYEFDGLDWDWEYP 121  
83 LKQNPFLKTLAIGWNFGTQFTDMVAANNRQTFVNSAIRFLKRYFDFGLDWDWEYP 142  
122 GSRGSPQDKHLFTVLVOEMREAFQEAQKINPKRLMVTAAVAAGISNTQSGYEIPQLSQ 181  
143 GSGQSPAVDKERTFTLVQDLANAFQEAQTSGRKRLLSAAVPAQOTYVDAGYEVKIAQ 202  
182 YLDYIHVMYDLHSGNEGTGENSPLYKPYPTDTSNAYLNVDYVMYKDNKGAPEKLIV 241  
203 NLDVFNLMAYDFHGSWEKVTGHNPSPLYKRYEESGAASLNVDVAQVQMLQKGPFPASKLIL 262  
242 GFPTYGHNFTLSPNTGICAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDAPO 301  
263 GMYTGRSFTLASSDTRVGAPATGSGTGPFTKEGMLAYEVCW--KGATKQRIQDQ 320  
302 EYVYAGVGNWVGYDNKSPDIIKAQWLKHNKFGAMVWAILDDFTGFCNQCKFPLIST 361  
321 KVPYIFRDQNWGVDFDVSFKTKVSYLKKQGLGCAVMWALDDDFAGFSCNQGRYPLIQ 380  
362 LKAL-----GLQASCTAPAQPIEPIITAAPSGSGNGSGSGSGSGGSCGFCVAVRAN 414  
381 LRQELSLPYLPSGTPELEVPKQPQSEP-----EHGFPSPGQDTFCQKAD 425

QY 415 GLYPVANNRNFHCVNGVTYQNCQAGLVFDTSCDCNNW 454  
DB 426 GLYPNRRSSPYSCAAGRLFQSCPTGLVFNCKCCTW 465

## RESULT 9

AAE00432  
ID AAE00432 standard; Protein; 466 AA.

XX AC AAE00432;

XX DT 19-JUN-2001 (first entry)

XX Human chitinase protein from clone pMO-218.

XX Human; antifungal; chitinase; immunoglobulin; Ig; therapy;

KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;

KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;

KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;

XX clone pMO-218.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..21

FT Protein /label= Signal\_peptide

FT Domain /label= Human\_mature\_chitinase\_protein

FT Region /label= Chitin\_binding\_domain

FT Note= "Region with triacetylchitotriose hydrolysing activity"

XX WO200123430-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26960.

XX 30-SEP-1999; 99US-0409918.

XX (ICOS-) ICOS CORP.

XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;

XX WPI; 2001-266141/27.

XX N-PSDB; AAD03759.

XX Novel chitinase immunoglobulin fusion product, useful for treating fungal infections and reducing the amount of a non-chitinase antifungal agent needed for the treatment

XX Claim 1; Page 32-33; 39pp; English.

XX The present invention relates to a chitinase immunoglobulin (Ig) fusion product, comprising a human chitinase fused to at least a portion of an immunoglobulin chain. The fusion product is useful for treating fungal infections (mycoses) such as candidiasis, aspergillosis, blastomycosis, coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis, cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses. The fusion protein is useful for reducing the amount of non-chitinase antifungal agent needed to exert an antifungal activity. The fusion protein is also useful for preparing a medicament for the prophylactic or therapeutic treatment of fungal infections. Chitinase immunoglobulin fusion product has unexpectedly improved serum half-life and formulation properties.

XX The present sequence is human chitinase protein from clone pMO-218. Chitinase enzyme degrades chitin which is a homopolymer of beta-(1,4)-linked N-acetylglucosamine residues.

XX Sequence 466 AA;



Query Match	52.8%;	Score 1323;	DB 22;	Length 466;
Best Local Similarity	52.2%;	Pred. No. 1e-111;		
Matches 240;	Conservative 70;	Mismatches 136;	Indels 24;	Gaps 3;
QY	2	QLTCYFTNWAQYRPGCLGRFMPDNDPCLCCTHLIIYAFAGRONNETTTTETWMDVTLQAFNG	61	
Db	23	KLVCYFTNWAQYRPGCLGRFMPDNDPCLCCTHLIIYAFAGMTNHQSLSTWENDETLYOEFNG	82	
QY	62	LKNKNSQLKTLIIAGGNWFGTAPFTAMVSTPENQTFITTSVIKFLROYEFDGLDFWEYP	121	
Db	83	LKQMPKLTLLAIGGNWFGTQKFTDMVATANNQTFVNSAIRFLRYKSPDGLDLDWEYP	142	
QY	122	SGRSGPPQDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAGISNIOGSEIPELSQ	181	
Db	143	GSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIAQ	202	
QY	182	YLDYIHVMTYDLHGSWEYTGENSEPLYKYPTDTSNAYLVNDYVMYWKONGAPAEKLIIV	241	
Db	203	NLDFVNLMAYPDHGSWEKVTGHNSPLYKQESGAAASLVNDAVQOQLQGTGTPASKLIL	262	
QY	242	GPFTYGHNFILSNPNTGIGAPTSGAGPAGYAKESGIWAYYEICTFLKNGATQGDWAPQ	301	
Db	263	GMPTVGRSFTLASSDTRVGPATGSGTGPFTKEGMLAYEYVCSW--KGATKQRIQDQ	320	
QY	302	EVPYAYQGNVWGYDNIKSFDIKQWLKHNNKFGGAMVWALDLDFTGFCNQKGFELIST	361	
Db	321	KVPYIFRDNQWGFDDVESFKTKVSLKQKGLGAMVWALDLDDBFAGFSCNQGRYPLIQT	380	
QY	362	LKAL-----GLQASCTAPAPIEPITAPSGSGNGSGSSSSGSGSGSGFCAVRAN	414	
Db	381	LRQELSLPYLPSGTPELVPKPGQSEP-----EHPGSPQDFTFCQKAD	425	
QY	415	GLYPVANNRNFHWCVNGVTYQONCOAGLVFDTSCDCNW	454	
Db	426	GLYPNPRSSSFYSCAAGRLFQOCSCTGLVFSNCKCCTW	465	
RESULT 10				
ID	ABB76291			
XX	ABB76291 standard; Protein; 466 AA.			
AC	ABB76291;			
DT	12-AUG-2002 (first entry)			
XX	Human chitinase.			
DE				
XX	Chitinase; enzyme; human; fungicide; antifungal; infection;			
KW	paracoccidioidomycosis; coccidioidomycosis; blastomycosis;			
KW	chromoblastomycosis; histoplasmosis; cryptococcosis;			
KW	chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;			
OS	Pneumocystis.			
XX	Homo sapiens.			
XX				
FT	Key	Location/Qualifiers		
FT	Peptide	1..21		
FT	Protein	/label= Signal_peptide		
FT		22..466		
FT		/label= Mature_protein		
XX	US6372212-B1.			
XX				
PD	16-APR-2002.			
XX				
PF	16-JUN-1997;	97US-0877599.		
XX				
PR	14-JUN-1996;	96US-0663618.		
XX				
PA	(ICOS-) ICOS CORP.			
XX				
Gray PW;				

```

RESULT 11
AAW40260
ID AAW40260 standard; Protein; 466 AA.
XX
XX AAW40260;
AC
XX
XX 15-JUN-1998 (first entry)
DT
XX
XX Human chitinase protein from clone MO-13B.
DE
XX
XX Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix.
XX
XX Homo sapiens.
OS
XX
XX
XX Key Location/Qualifiers
FH Peptide 1..21
FT Protein /label= signal
FT Protein 22..466
FT Protein /label= chitinase
FT Protein /note= "from clone MO-13B"
XX
XX
XX MO9747752-A1.
PN
XX
XX 18-DEC-1997.
XX
XX 16-JUN-1997; 97WO-US10460.
XX
XX 14-JUN-1996; 96US-0663618.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW;
XX
XX WPI; 1998-052316/05.
DR N-PSDB; AAV10436.
XX
XX Nucleic acids encoding human chitinase - useful as antifungal
PT agents, especially in combination with other antifungals
XX
XX Claim 7; Page 44-45; 63pp; English.
XX
XX This sequence represents a novel human chitinase isolated from clone
CC MO-13B. Chitinases are useful for treating or preventing fungal infection
CC and as immunogens for generating antibodies which are used to purify,
CC detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.
CC The nucleic acid sequence of the chitinase is also useful as a probe to
CC identify and isolate genomic DNA encoding chitinases or similar proteins,
CC or cells expressing them or to generate transgenic ('knockout') rodents.
CC It can also be used in hybridisation assays and to detect genetic
CC alterations in the chitinase gene related to disease. Agents that inhibit
CC this protein may be useful in treatment of Gaucher's disease and
CC rheumatoid arthritis, where overexpression of the protein can damage
CC the extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,
CC and thus of their side effects.
XX
XX Sequence 466 AA;
XX
XX Query Match 52.6%; Score 1317; DB 19; Length 466;
XX Best Local Similarity 52.0%; Pred. No. 3.6e-111;
XX Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;
XX
XX 2 QLTCTYTNWAQRPGLGRFPDNDICLCTHLYIYAFAGQNNETIWNVDLYQAFNG 61
XX : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 23 KLVCTYTNWAQYRQGRARLPKDLPSLCTHLYIYAFAGMTNQLSTWENDELYQEFNG 82
XX : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 62 LKNKNSQLKTLAIGGNWFTAPFTAMVSPENRQFTFVTKFLQRYEFDGLDFDWEYP 121
XX : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 83 LKQNFPLKTLAIGGNWSTQFTDQVATANNRQIFVNSAIRFLKRYEFDGLDWEYP 142
XX : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 122 GSRGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAGISNTQSGYEIPQLSQ 181
XX : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 143 GSQGS PAVDKERFTTLVQDLANAFQCEAQTSGKERLLLSAAVPAGQTYVDAGYVDKIAQ 202
Qy 182 YLDYIHWMTYDLHGSWEQYTGNSPLYKYPTDTGSNAVILNDVYVMYKDKNGCAPAEKLV 241
Db 203 NLDYFNLMAIDYFHGSWEKVTGHSPLYKQREESGAAASLNDVDAVQOQWLQKGT PASKLIL 262
Qy 242 GPPTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGWDAPO 301
Db 263 GNPVTGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYVEVCSW--KGATKQRIQDO 320
Qy 302 EYPIAYQGVVWYGYDNIKSFDIKAWLKHKNKFGGAMVMAIDLDFTGTFCNQGKPEPLIST 361
Db 321 KVPYIFRDNQWVGFDDVESFTKVSYLKQKGLGGAMVMAIDLDLDDPAGFSNCGRYPLIQ 380
Qy 362 LKKAL-----GLQASCTAPAOPIETAPSGSGSGSGSGSGSGSGSGSGSGSGSGSG 414
Db 381 LRQELSLPYLPSTGTPLEVPKGPQSEP-----EHPSPGQDTFCQGRAD 425
Qy 415 GLYPVANNRNFHVCVNGVTYQNCQAGLVFDTSCDCCNW 454
Db 426 GLYPNPRERSFYSACAGRLFOQSCPTGLVFSNSCKCCTW 465

RESULT 12
AAW42426
ID AAW42426 standard; Protein; 466 AA.
XX
XX AAW42426;
AC
XX
XX 10-DEC-1999 (first entry)
DT
XX
XX MO-13B clone of human Chitinase, amino acid sequence.
DE
XX
XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin-binding; allele; vector;
KW truncated protein; chitin binding region.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..21
FT Protein /label= Signal_peptide
FT Protein 22..466
FT Protein /label= Mature_protein
XX
XX WO9946390-A1.
XX
XX 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US05343.
XX
XX 12-MAR-1998; 98US-0039198.
XX (ICOS-) ICOS CORP.
XX
XX Gray PW, Tjoelker LW;
XX
XX WPI; 1999-551417/46.
DR N-PSDB; AA221848.
XX
XX Novel chitin-binding fragments of human chitinase used to treat fungal
PT infections in animals
XX
XX Example 1; Page 62-64; 83pp; English.
XX
XX This is the amino acid sequence of an allelic form of the human
CC chitinase enzyme, which is capable of degrading Chitin (a linear
CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).
CC The sequence also includes a Chitin binding peptide region, which lacks
CC any chitinase activity, but which can be used to target anti-fungal
CC agents to chitin containing fungal cells.
CC Chitinase fragments can be used to screen for proteins or other

```

CC molecules that specifically bind to the chitin-binding domain of human  
 CC chitinase or that modulate its activity. These compounds are useful for  
 CC immunization, as well as for purifying chitinase, as well as for  
 CC detection and quantification of chitinase. Polynucleotide fragments of  
 CC the invention are useful as a source of probes and primers, and to  
 CC express the proteins recombinantly. The chitinase fragments, when  
 CC conjugated to antifungal compounds, are used to treat animals,  
 CC especially humans, infected with chitin-containing parasites such as  
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
 CC sporotrichosis, and dermatophytoses.  
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
 CC protein for treating infections, especially fungal infections, is  
 CC problematic. In view of the increasing incidents of life-threatening  
 CC fungal infection in e.g. immunocompromised individuals, there exists a  
 CC need for identifying new compounds for treating fungal infection. The  
 CC chitin-binding fragments of the present invention provide this need.  
 XX  
 SQ Sequence 466 AA;

Query Match 52.6%; Score 1317; DB 20; Length 466;  
 Best Local Similarity 52.0%; Pred. No. 3.6e-111;  
 Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;  
 QY 2 QLTCTFTWNAQYRPGLGFRMPNDIPCLCTHLIYAFAGRNNEITTEWNDVTLQAENG 61  
 DB 23 KLVCTFTWNAQYRQGEARFLPKDLPCLCTHLIYAFAGMTNQLSTTEWDELTYQEFNG 82  
 QY 62 LKKNKSQLKTLIAIGWNFGTAPFTAMVSTPENROTFITSVIKLRQYEFDFGLDFWEYP 121  
 DB 83 LKKNPKLKTLLAIGWNFSTQKFTDMVATANNRQTFVNSAIRFLKYSFDFGLDWEYP 142  
 QY 122 GSRGSPDQKHFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNIQSYEIPQLSQ 181  
 DB 143 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQTYVDAGYEVDKIAQ 202  
 QY 182 YLDYTHVMTYDLHGSWEGYTGNSPLYKYPTDTGSMAYLVNDYVWYWKDNGAPAEKLTIV 241  
 DB 203 NLDVFNLMAYDFHGSWEKVTGHSNPLYKQESGAASLVNDAVQOVLQKTPASKLIL 262  
 QY 242 GPPTGHNFLNPNSTNGIAPTSAGPAGPVAKSGIWAYEYICFTFLKNGATQGWDAPO 301  
 DB 263 GMPTYGRSFTLASSSDTRVAGPATSQGTGPTKGGMLAYEVCWS--KGATKQRIQDQ 320  
 QY 302 EYPYAYQGNVWGYDNIKSFDIAKWLKHNKFGGAMWALDIDFTGTCNKGKPLIST 361  
 DB 321 KVPYIFRDQWGVDFDVFSTKVSYLKQKGGAMWALDIDDFAGFCNQGRYPLIQT 380  
 QY 362 LKAL-----GLQASACTAPAQPIETITAPSGSGNGSGSGSGSGSGSGGFCVAVRAN 414  
 DB 381 LKQELSLPVLPSGTPLEVPKQPSF-----EKGPSFGQDTFCQKAD 425  
 QY 415 GLYPVANNRNAPWHCVNGVYQONQOAGLVFTSDCCNW 454  
 DB 426 GLYPNPRRSFSCAAGRLFQSCPTGLVFNCSCKCTW 465

## RESULT 13

AAE00433

XX

AC

XX

DT

XX

DE

XX

KW

KW

KW

KW

XX Homo sapiens.  
 OS  
 XX  
 FH Key  
 FT Peptide  
 FT Protein  
 FT Domain  
 FT Region  
 FT  
 FT Location/Qualifiers  
 /label= Signal\_peptide  
 1..21  
 22..466  
 /label= Human\_mature\_chitinase\_protein  
 418..466  
 /label= Chitin\_binding\_domain  
 1..373  
 /note= "Region with triacetylchitotriose  
 hydrolysing activity"

WO200123430-A2.

05-APR-2001.

28-SEP-2000; 2000WO-US26960.

30-SEP-1999; 99US-0409918.

(ICOS-) ICOS CORP.

Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;

WPI: 2001-266141/27.

N-PSDB; AAD03760.

XX Novel chitinase immunoglobulin fusion product, useful for treating  
 PT fungal infections and reducing the amount of a non-chitinase antifungal  
 PT agent needed for the treatment -  
 XX

Claim 1; Page 36-38; 39pp; English.

XX The present invention relates to a chitinase immunoglobulin (Ig) fusion  
 CC product, comprising a human chitinase fused to at least a portion of an  
 CC immunoglobulin chain. The fusion product is useful for treating fungal  
 CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,  
 CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
 CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.  
 CC The fusion protein is useful for reducing the amount of non-chitinase  
 CC antifungal agent needed to exert an antifungal activity. The fusion  
 CC protein is also useful for preparing  
 CC a medicament for the prophylactic or therapeutic treatment of fungal  
 CC infections. Chitinase immunoglobulin fusion product has unexpectedly  
 CC improved serum half-life and formulation properties.  
 CC The present sequence is human chitinase protein from clone pMO-13B.  
 CC Chitinase enzyme degrades chitin which is a homopolymer of  
 CC beta-(1,4)-linked N-acetylglucosamine residues.

XX Sequence 466 AA;

Query Match 52.6%; Score 1317; DB 22; Length 466;

Best Local Similarity 52.0%; Pred. No. 3.6e-111;

Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;

QY 2 QLTCTFTWNAQYRPGLGFRMPNDIPCLCTHLIYAFAGRNNEITTEWNDVTLQAENG 61

DB 23 KLVCTFTWNAQYRQGEARFLPKDLPCLCTHLIYAFAGMTNQLSTTEWDELTYQEFNG 82

QY 62 LKKNKSQLKTLIAIGWNFGTAPFTAMVSTPENROTFITSVIKLRQYEFDFGLDFWEYP 121

DB 83 LKKNPKLKTLLAIGWNFSTQKFTDMVATANNRQTFVNSAIRFLKYSFDFGLDWEYP 142

QY 122 GSRGSPDQKHFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNIQSYEIPQLSQ 181

DB 143 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQTYVDAGYEVDKIAQ 202

QY 182 YLDYTHVMTYDLHGSWEGYTGNSPLYKYPTDTGSMAYLVNDYVWYWKDNGAPAEKLTIV 241

DB 203 NLDVFNLMAYDFHGSWEKVTGHSNPLYKQESGAASLVNDAVQOVLQKTPASKLIL 262

242 GPPTYGHNFIILNSNTGIGAPTSAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301  
263 GNPYGRSFTLASSDTRVGAATGSGTGPFTKEGMLAYEVCWS--KGATKQRIQDO 320  
302 EYPYAYQGNVWGYDNIKSFIDIAQWLKHNKFGGMVAIDLDFTGTCNCKGKPLIST 361  
321 KVPYIFRDNQWGFDDVESFKTVSYLKQKGLGGAMVWALDLDFAFGSCNCRYPLOT 380  
362 LKAL-----GLQASCTAPAPQIEPIITAAPSGSGSGSGSGSGSGSGGFCAVRAN 414  
381 LRQELSLPVLPSGTPELEVPKQPSEP-----EHGPSQDQTFQCGKAD 425  
415 GLYPVANNNFHWCVNGVYQONCOAGLVFTSDCCCNW 454  
426 GLYPNPRSSFYSCAAGRLFQOSCTGLVFSNCKCCTW 465

RESULT 14  
ID ABB76292 standard; Protein; 466 AA.  
XX ABB76292;  
XX 12-AUG-2002 (first entry)  
DT Human chitinase.  
DE  
XX Chitinase; enzyme; human; fungicide; antifungal; infection;  
KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;  
KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
KW Pneumocystis.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..21 /label= Signal\_peptide  
FT Protein 22..466 /label= Mature\_protein  
FT  
XX US6372212-B1.  
XX  
XX 16-APR-2002.  
XX 16-JUN-1997; 97US-0877599.  
XX 14-JUN-1996; 96US-0663618.  
XX (ICOS-) ICOS CORP.  
XX Gray PW;  
XX WPI; 2002-442449/47.  
XX N-PSDB; ABL57381.  
XX  
XX Co-administering chitinase to improve the effectiveness of fungicidal  
PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis  
XX  
XX Example 1; Column 33-36; 26pp; English.  
XX  
XX The present sequence is the protein sequence of human chitinase  
CC as predicted from isolated cDNA clone MO-13B (see ABL57378). It  
CC differs in only 1 amino acid residue from the chitinase sequence  
CC (see ABB76291) deduced from a second cDNA clone, having serine  
CC at position 81 of the mature protein. Northern blots showed  
CC highest chitinase gene expression in lung and ovary tissues.  
CC Expression in lung is consistent with a protective role against  
CC pathogenic organisms that contain chitin. The invention provides  
CC human chitinase polynucleotides and polypeptides, and materials and  
CC methods for the recombinant production of human chitinase products,  
CC which are expected to be useful as products for treating fungal .

infections or for the development of such products. Human  
chitinase has a synergistic effect on the actions of other  
fungicides. It can be administered to improve the antifungal  
activity of a non-chitinase antifungal agent, especially  
amphotericin B or itraconazole, in the treatment of a fungal  
infection such as candidiasis, aspergillosis, coccidioidomycosis,  
blastomycosis, paracoccidioidomycosis, histoplasmosis,  
cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,  
dermatophytoses and pneumocystis infections (all claimed). In  
particular, the fungal infection involves Candida, Aspergillus  
and/or Cryptococcus spp., whose growth is not effectively  
inhibited by contact with human chitinase alone.

Query Match 52.6%; Score 1317; DB 23; Length 466;  
Best Local Similarity 52.0%; Pred. No. 3.6e-111; Indels 24; Gaps 3;  
Matches 239; Conservative 70; Mismatches 121;

QY 2 QLTCTFTNNAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIEMNDVTLYQAFNG 61  
DB 23 KLVCTFTNNAQYRQGEARFLPKLDPSLCTHLIYAFAGMTNHQSTTEWDETLYQEFNG 82  
QY 62 LKNKNSQLKTLIAIGWNFGTAPFTAMYSTPENROTFTTSVTKFLROYEFDGLDFDWEYP 121  
DB 83 LKKMNPCLKTLIAIGWNFSTQKFTDMVATANNROTFFVNSAIRFLKYSFDFGLDLDWEYP 142  
QY 122 GSRGSPPODKHLFTVLVQEMREAPQEAQKINPKRLMTAAVAAGISNIQSGYEIPOLSO 181  
DB 143 GSQGSAPVDKERTFTLVQDLANAFQOEATQSKERLLLSAAVPAGQTYVDAGYEVDKIAQ 202  
QY 182 YLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSAVLNVDYVMNYKNDGAPAEKLIIV 241  
DB 203 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOQLKGTASKLIL 262  
QY 242 GPPTYGHNFIILNSNTGIGAPTSAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301  
DB 263 GNPYGRSFTLASSDTRVGAATGSGTGPFTKEGMLAYEVCWS--KGATKQRIQDO 320  
QY 302 EYPYAYQGNVWGYDNIKSFIDIAQWLKHNKFGGMVAIDLDFTGTCNCKGKPLIST 361  
DB 321 KVPYIFRDNQWGFDDVESFKTVSYLKQKGLGGAMVWALDLDFAFGSCNCRYPLOT 380  
QY 362 LKAL-----GLQASCTAPAPQIEPIITAAPSGSGSGSGSGSGSGSGGFCAVRAN 414  
DB 381 LRQELSLPVLPSGTPELEVPKQPSEP-----EHGPSQDQTFQCGKAD 425  
QY 415 GLYPVANNNFHWCVNGVYQONCOAGLVFTSDCCCNW 454  
DB 426 GLYPNPRSSFYSCAAGRLFQOSCTGLVFSNCKCCTW 465

RESULT 15  
AAW31498  
ID AAW31498 standard; Protein; 466 AA.  
XX AAW31498;  
XX 27-APR-1998 (first entry)  
XX Human chitotriosidase variant.  
XX  
XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;  
KW rheumatoid arthritis; atherosclerosis; human.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FT Misc-difference 102 /note= "encoded by RGC"  
XX W09736917-A1.  
XX



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 21:01:56 ; Search time 17.5 Seconds  
(without alignments)  
2850.818 Million cell updates/sec

Title: US-10-004-219B-14

Perfect score: 2506

Sequence: 1 YQLTCYFTNWAQYRGLGRF.....QQNCQAGLVFTSDCCNWA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pap.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pap.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pap.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pap.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pap.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pap.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pap.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pap.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2506	100.0	455	9	US-10-004-219B-14
2	2506	100.0	476	9	US-10-004-219B-14
3	2089.5	83.4	452	9	US-10-004-219B-9
4	2089.5	83.4	473	9	US-10-004-219B-4
5	1323	52.8	445	9	US-10-004-219B-10
6	976.5	39.0	383	10	US-10-097-340-45
7	968.5	38.6	383	10	US-09-459-749D-17
8	711.5	28.4	490	9	US-10-218-743-41
9	711.5	28.4	509	9	US-10-218-743-35
10	711.5	28.4	509	9	US-10-218-743-38
11	708.5	28.3	536	9	US-10-218-743-21
12	708.5	28.3	555	9	US-10-218-743-15
13	708.5	28.3	555	9	US-10-218-743-18
14	472	18.8	371	9	US-09-923-844B-2
15	285.5	11.4	376	10	US-09-748-033-3
16	196	7.4	170	9	US-10-218-743-44
17	104.5	4.2	711	9	US-10-071-485-90
18	104	4.2	309	9	US-10-304-928-10
19	98.5	3.9	1385	10	US-09-738-363-2

20	98	3.9	750	9	US-10-046-433-2	Sequence 2, Appli
21	98	3.9	1001	9	US-10-046-433-40	Sequence 40, Appl
22	98	3.9	1013	9	US-10-028-072-38	Sequence 38, Appl
23	98	3.9	1013	9	US-10-121-049-38	Sequence 38, Appl
24	98	3.9	1013	9	US-10-123-004-38	Sequence 38, Appl
25	98	3.9	1013	9	US-10-140-470-38	Sequence 38, Appl
26	98	3.9	1013	9	US-10-175-746-38	Sequence 38, Appl
27	98	3.9	1013	9	US-10-176-918-38	Sequence 38, Appl
28	98	3.9	1013	9	US-10-176-921-38	Sequence 38, Appl
29	98	3.9	1013	9	US-10-137-865-38	Sequence 38, Appl
30	98	3.9	1013	9	US-10-140-474-38	Sequence 38, Appl
31	98	3.9	1013	9	US-10-142-431-38	Sequence 38, Appl
32	98	3.9	1013	9	US-10-143-114-38	Sequence 38, Appl
33	98	3.9	1013	9	US-10-140-002-38	Sequence 38, Appl
34	98	3.9	1013	9	US-10-142-419-38	Sequence 38, Appl
35	98	3.9	1013	9	US-10-123-262-38	Sequence 38, Appl
36	98	3.9	1013	9	US-10-142-423-38	Sequence 38, Appl
37	98	3.9	1013	9	US-10-121-050-38	Sequence 38, Appl
38	98	3.9	1013	9	US-10-141-755-38	Sequence 38, Appl
39	98	3.9	1013	9	US-10-143-032-38	Sequence 38, Appl
40	98	3.9	1013	9	US-10-123-108-38	Sequence 38, Appl
41	98	3.9	1013	9	US-10-123-236-38	Sequence 38, Appl
42	98	3.9	1013	9	US-10-123-261-38	Sequence 38, Appl
43	98	3.9	1013	9	US-10-140-921-38	Sequence 38, Appl
44	98	3.9	1013	9	US-10-140-928-38	Sequence 38, Appl
45	98	3.9	1013	9	US-10-121-045-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1  
US-10-004-219B-14  
; Sequence 14, Application US/10004219B  
; Publication No. US20030087414A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; FILE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)..(455)  
; OTHER INFORMATION: /note="Human AMCase"  
US-10-004-219B-14

Query Match	100.0%	Score 2506;	DB 9;	Length 455;
Best Local Similarity	100.0%;	Pred. No. 6e-204;	Mismatches 0;	Indels 0; Gaps 0;
Matches 455;	Conservative 0;			
Qy	1	YQLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRQNNETTIEWNDVTLYQAFN	60	
Db	1	YQLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRQNNETTIEWNDVTLYQAFN	60	
Qy	61	GLKKNKSQLKTLAIGGNWFGTAPMTAVSTPENRQTITSVIKFLRQYFDFGLDFDWEY	120	
Db	61	GLKKNKSQLKTLAIGGNWFGTAPMTAVSTPENRQTITSVIKFLRQYFDFGLDFDWEY	120	
Qy	121	PGSRGSPQDKHLFTVLVOEMREAFQEAQKINKPRLMVTAAVAGISNIQSGYEIPQLS	180	
Db	121	PGSRGSPQDKHLFTVLVOEMREAFQEAQKINKPRLMVTAAVAGISNIQSGYEIPQLS	180	

### RESULT 3

```

: APPLICANT: Aerts, Johannes M.F.G.
: APPLICANT: Boot, Rolf G.
: TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
: TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
: TITLE OF INVENTION: which mucus is involved or infection diseases
: FILE REFERENCE: 2183-5136US
: CURRENT APPLICATION NUMBER: US/10/004,219B
: CURRENT FILING DATE: 2001-11-02
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 9
: LENGTH: 452

```

```

? LENGTH: 452
? TYPE: PRT
? ORGANISM: Mus musculus
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (1)..(452)
? OTHER INFORMATION: /note="Mouse AMCase"
US-10-004-219B-9

```

	Query Match	83.4%	Score 2089.5	DS: 9	Length 452
	Best Local Similarity	81.3%	Pred. No. 1.1e-168		
	Matches 369	Conservative 40	Mismatches 42	Indels 3	Gaps 1
QY	1	YQLTCYFTNWAQYRPGLGRRMPONIDPCLCTHLIYAFAGQNNRITTEIEMNDVTLYQAFN	60		
Db	1	YNLICVFTNWAQYRPGLGSGFKPDINPCLCTHLIYAFAGMONNRIITTEIEMNDVTLYKAFN	60		
QV	61	GLKNKNSOLKTLAIGGWNFGTAPPTAMVSTPNRQTFITSVIKRLOYEDGDLDFDWEY	120		

61	DLNRNSKLTLLAIGWNFGTAPFTMTVSTSNRQTFITSVIKFLKQYGFDDGLDWEY	120
Db		
121	PGSRGSPPODKHLFTVLVQEMRFAFEQAEKQINKPLRMVTTAAVAGISNIOGSEYIPELS	180
Qy		
121	PGSRGSPPODKHLFTVLVQEMRFAFEQAEIESTNRPLRMVTTAAVAGGISNIOGSEYIPELS	180
Db		
181	QYLDYTHVMYTDYDLHGSGWEGYTGNSPLYKPYPTDTGTSNAYLNVDYVMYWNKONGAPAEKLI	240
Qy		

[illegible][illegible]



## RESULT 4

US-10-004-219B-4  
; Sequence 4, Application US/10004219B  
; Publication No. US20030087414A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 473  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: mouse AWCae  
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence  
US-10-004-219B-4

Query Match	83.4%;	Score	2089.5;	DB 9;	Length	473;			
Best Local Similarity	81.3%;	Pred. No.	1.2e-168;						
Matches	369;	Conservative	40;	Mismatches	42;	Indels	3;	Gaps	1;

QY 1 YQTCYFTNWAQYRPGLGFMFNDIPCLCTHLIYAFAGQNNETIIEWNDVTLYQAFN 60  
DB 22 YNLICVFTNWAQYRPGLGSKFDPDINPCILCTHLIYAFAGQNNETIIEWNDVTLKAFN 81  
QY 61 GLKKNLSQKTLTLLAIGGWNFGTAPFTAMYSTENRQTFITSVIKFLROYEFDGLDWEY 120  
DB 82 DLKRNRSKLTLLAIGGWNFGTAPFTMTVSTQNROTFTITSVWIKFLROYEFDGLDWEY 141  
QY 121 PGSRSGPPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAAGISNIQSGYEIPQLS 180  
DB 142 PGSRSGPPQDKHLFTVLVQEMREAFQEALESNRPRLMTAAVAGISNIQAGYEIPELS 201  
QY 181 QYLDYTHVMTYDLHGSWEGYTGENSPLYKYPTDTGNSNAYLNDVYVNNYKNGAPAEKLI 240  
DB 202 KYLDFTHVMTYDLHGSWEGYTGENSPLYKYPTETGNSNAYLNDVYVNNYKNGAPAEKLI 261  
QY 241 VGFTYGHNFILNSPNTGIGAPTSAGAPYAKESGWAYYEICTFLKNGATQGWADAP 300  
DB 262 VGFTYGHNFILNSPNTGIGAPTSAGAPYAKESGWAYYEICTFLKNGATQGWADAS 321  
QY 301 QEVYAYQGVNVMVGYDNIKSFDIKAQWLKHNKFGGAMVAIDLDDFTGTCNQGKPEPLIS 360  
DB 322 QEVYAYKANEMLYDNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQKPEPLTS 381  
QY 361 TLKALGLQASCTAPAPQPIETAPSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 420  
DB 382 TLKALGISTEGCTAPDVPSEPTTTP--GSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 438  
QY 421 NNRNAPFWHCVNGVTYQONCOAGLVFDTSCDCCNW 454  
DB 439 DORNAPWQINGITYQOHCOAGLVFDTSCNCCNW 472

## RESULT 5

US-10-004-219B-10  
; Sequence 10, Application US/10004219B  
; Publication No. US20030087414A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases

FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 445  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)...(445)  
; OTHER INFORMATION: /note="Human chitotriosidase"  
US-10-004-219B-10

Query Match	52.8%;	Score	1323;	DB 9;	Length	445;			
Best Local Similarity	52.2%;	Pred. No.	8.5e-104;						
Matches	240;	Conservative	70;	Mismatches	126;	Indels	24;	Gaps	3;

QY 2 QTCYFTNWAQYRPGLGFMFNDIPCLCTHLIYAFAGQNNETIIEWNDVTLYQAFNG 61  
DB 2 KLVCFYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFNG 61  
QY 62 LKNKNSQLKTLTLLAIGGWNFGTAPFTAMYSTENRQTFITSVIKFLROYEFDGLDWEY 121  
DB 62 LKNKNSQLKTLTLLAIGGWNFGTAPFTAMYSTENRQTFITSVIKFLROYEFDGLDWEY 121  
QY 122 GSRSGPPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAAGISNIQSGYEIPQLS 181  
DB 122 GSGSPAVDKERFTVLVQDLANAFQEAQTSCKERLLLSAAVPAQYTVYDAGYVDKIAQ 181  
QY 182 YLDYTHVMTYDLHGSWEGYTGENSPLYKYPTDTGNSNAYLNDVYVNNYKNGAPAEKLI 241  
DB 182 NLDVNLMAVDYFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQWLQKGTASKLIL 241  
QY 242 GPTVYGHNFILNSPNTGIGAPTSAGAPYAKESGWAYYEICTFLKNGATQGWADAP 301  
DB 242 GMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGLAYYEVCWS--KGATKQRIQDQ 299  
QY 302 EYVYAYQGVNVMVGYDNIKSFDIKAQWLKHNKFGGAMVAIDLDDFTGTCNQGKPEPLIS 361  
DB 300 KVPYIFRDNQWVGFDVDFESFTKYSYLKQKGLGGAMVAIDLDDFTGTCNQGKPEPLIS 359  
QY 362 LKAL-----GLQASCTAPAPQPIETAPSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 414  
DB 360 LKQELSLPLPSGTPELEVPRKGPSEP-----EHGSPSGQDTFCQKAD 404  
QY 415 GLYPVANNRNPFWHCVNGVTYQONCOAGLVFDTSCDCCNW 454  
DB 405 GLYPNPRSSPFSYCAAGRLFOQSCPTGLVFSNCKCCTW 444

## RESULT 6

US-10-097-340-45  
; Sequence 45, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Marjula GANNAVARAPU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAMATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT

TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
Assessment, Prevention, and Therapy of Ovarian Cancer  
FILE REFERENCE: MRI-030  
CURRENT APPLICATION NUMBER: US/10/097,340  
CURRENT FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: 60/276,025  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/325,149  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/276,026  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/324,967  
PRIOR FILING DATE: 2001/09/26  
PRIOR APPLICATION NUMBER: 60/311,732  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: 60/325,102  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/323,580  
PRIOR FILING DATE: 2001-09-19  
NUMBER OF SEQ ID NOS: 363  
SOFTWARE: Fast-Seq for Windows Version 4.0  
SEQ ID NO 45  
\* TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-097-340-45

Query Match 39.0%; Score 976.5; DB 9; Length 383;  
Best Local Similarity 51.2%; Pred. No. 1.5e-74;  
Matches 188; Conservative 62; Mismatches 108; Indels 9; Gaps 5;  
QY 1 YQTCYFTNWAQYRGLGRMPDNIDPCLCTHLYAFAGRQNNETITIEWNDVTLQAFN 60  
DB 22 YKLVCTYTSWSQYREGDSCFPDAPLDRFLCTHLYIFSFANISNNEIDTLEWNDVTLQMLN 81  
QY 61 GLKKNKSQLKTLAIGGNWFGTAPMTAVSTPENRQTFTITSVIKFLRQYEFDFGLDPWEY 120  
DB 82 TLKRNPNLKTLLSVGWNFGSQRFKSIASNTQSRRTFKSVPPFLRTHGFDGLDLAWLY 141  
QY 121 PGSRGSPQDKHLFTVLVQEMREAFQEAQKINKPLMTAAVAAGISNIQSYEIPQLS 180  
DB 142 PGR-----DKQHFTLLIKEMKAEFTKEA-OPGKQLLSAALSAGKVTIDSSYDIKIS 195  
QY 181 QYLDYIHVMYDHLGWSWEGYTGNSPLKYPTDTGNSAYLVNDVYVMYKNDGAPAEKLI 240  
DB 196 QHLDIFSIMTYDFHGAWRTTGHHSPLFRQOEASDPRFSNTDYAVGYMLRGLGAPASKLV 255  
QY 241 VGPTTYGHNFLNSNTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDP 300  
DB 256 MGIPTFGRSFTLAS-SETGVGAPISGPIGRPTKEAGTLAYYEICDFLR-GATVHRTLG 313  
QY 301 QYVPYAYQGNVWGYDNIKSFDIKAQWLKHNKFGAMVWALDLDGTFGFCNQG-KFPLI 359  
DB 314 QQVPYATKGNQWGYDDQESVKSQVLYKDRQLAGAMVWALDLDGTFGFCNQGLFFPLT 373  
QY 360 STLKAL 366  
DB 374 NAIKDAL 380

RESULT 7  
US-09-459-749D-17  
Sequence 17, Application US/09459749D  
Patent No. US20020136716A1  
GENERAL INFORMATION:  
APPLICANT: Millis, Albert J. T.  
TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration  
FILE REFERENCE: 0794.016A  
CURRENT APPLICATION NUMBER: US/09/459,749D  
CURRENT FILING DATE: 1999-12-10  
PRIOR APPLICATION NUMBER: 60/111,856  
PRIOR FILING DATE: 1998-12-11  
NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 17  
LENGTH: 383  
TYPE: PRT  
ORGANISM: Sus scrofa  
US-09-459-749D-17  
Query Match 38.6%; Score 968.5; DB 10; Length 383;  
Best Local Similarity 51.5%; Pred. No. 7.2e-74;  
Matches 189; Conservative 57; Mismatches 112; Indels 9; Gaps 5;  
QY 1 YQTCYFTNWAQYRGLGRMPDNIDPCLCTHLYAFAGRQNNETITIEWNDVTLQAFN 60  
DB 22 YKLVCTYTSWSQYREGDSCFPDAPLDRFLCTHLYIFSFANISNNEIDTLEWNDVTLQMLN 81  
QY 61 GLKKNKSQLKTLAIGGNWFGTAPMTAVSTPENRQTFTITSVIKFLRQYEFDFGLDPWEY 120  
DB 82 TLKRNPNLKTLLSVGWNFGSQRFKSIASNTQSRRTFKSVPPFLRTHGFDGLDLAWIS 141  
QY 121 PGSRGSPQDKHLFTVLVQEMREAFQEAQKINKPLMTAAVAAGISNIQSYEIPQLS 180  
DB 142 PGR-----DKQHFTLLIKEMKAEFTKEA-OPGKQLLSAALSAGKVTIDSSYDIKIS 195  
QY 181 QYLDYIHVMYDHLGWSWEGYTGNSPLKYPTDTGNSAYLVNDVYVMYKNDGAPAEKLI 240  
DB 196 QHLDIFSIMTYDFHGAWRTTGHHSPLFRQOEASDPRFSNTDYAVGYMLRGLGAPASKLV 255  
QY 241 VGPTTYGHNFLNSNTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDP 300  
DB 256 MGIPTFGRSFTLAS-SETGVGAPISGPIGRPTKEAGTLAYYEICDFLR-GATVHRTLG 313  
QY 301 QYVPYAYQGNVWGYDNIKSFDIKAQWLKHNKFGAMVWALDLDGTFGFCNQG-KFPLI 359  
DB 314 QQVPYATKGNQWGYDDQESVKSQVLYKDRQLAGAMVWALDLDGTFGFCNQGLFFPLT 373  
QY 360 STLKAL 366  
DB 374 NAIKDAL 380

RESULT 8  
US-10-218-743-41  
Sequence 41, Application US/10218743  
Publication No. US20030096779A1  
GENERAL INFORMATION:  
APPLICANT: McCall, Catherine A.  
APPLICANT: Hunter, Shirley Wu  
TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
TITLE OF INVENTION: AND USES THEREOF  
FILE REFERENCE: AL-2-C3  
CURRENT APPLICATION NUMBER: US/10/218,743  
CURRENT FILING DATE: 2002-08-13  
PRIOR APPLICATION NUMBER: US/09/292,225  
PRIOR FILING DATE: 1999-04-15  
PRIOR APPLICATION NUMBER: 60/098,909  
PRIOR FILING DATE: 1998-09-02  
PRIOR APPLICATION NUMBER: 60/085,295  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/098,565  
PRIOR FILING DATE: 1998-04-17  
PRIOR APPLICATION NUMBER: 09/062,013  
PRIOR FILING DATE: 1998-04-17  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 41  
LENGTH: 490  
TYPE: PRT  
ORGANISM: Dermatophagoides farinae  
US-10-218-743-41  
Query Match 28.4%; Score 711.5; DB 9; Length 490;  
Best Local Similarity 36.4%; Pred. No. 5.8e-52;

Matches 148; Conservative 74; Mismatches 152; Indels 33; Gaps 10;

QY 2 QLTCTTWAQYRPGLRMPNDIDPCLCTHLIYAFAGQNNETIENW----- 51  
Db 15 RIVCYVGTWSVYHK-VDPYTTIEDIDPFKCTHLMYGFADIDYKTYIQVDFPQDDNNHNSW 73

QY 52 DVTLYQAFNGLKNKNSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVKFLRQYEF 111  
Db 74 EKHGVERFNNLRKNPELTMTISLGGWEGSEKYSDMAANPTYRQVQSVDLDFLOEYKF 133

QY 112 DGLDFDWEYPGSR-GSPPODKHLFTVLVOEMREAFQEAQINKRPLMVTAAVAGISNI 170  
Db 134 DGLDLDWEYPGSRGKPKIDKQNYLTTLVRELKEAFEPFG-----YLLTAASVPGKDKI 186

QY 171 QSGYEIPOLSOYLDYIHVMTYDLHGSWEGYTGENSPLYKYPTDGTGS-NAYLNVDVYMNW 229  
Db 187 DVAYELKELNQLFDWMNVMTYDHGWNVFVGHNAPLYKRPDETDELHTYFNVTMTHY 246

QY 230 KONGAPAEKLVIGFPTYGHNFIILSNPSTGIGAPTSAGPAGPYAKESGIWAYEIC-TF 288  
Db 247 LNNGATRDKLNVGVPFYGYRAWSIEDRSKVKGDPKAGMSPPGFITGEGVLSYIELCQLF 306

QY 289 LKNGATQGWADAPQEVYAYQGNVWVGYNDIKSFQDKAQLKHNKFGGAMVAIDDDFTG 348  
Db 307 QKEEWHIQDEYNNAPYGYNDIKWGYDDLASISCKLAFKELGVSVMWISLENDDFKG 366

QY 349 TFCNKGKPELSTLKKALG-----LQASCTAPAPQIEPITAA 386  
Db 367 -HCGP-KYPLLKNVHNMINGDEKNSYECLLGPSTTPT-PTTPTSTS 410

## RESULT 9

US-10-218-743-35  
; Sequence 35, Application US/10218743  
; Publication No. US20030096779A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine A.  
; APPLICANT: Hunter, Shirley Wu  
; APPLICANT: Weber, Eric R.  
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
; FILE REFERENCE: AL-2-C3  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR FILING DATE: 1998-04-17  
; PRIOR FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 35  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Dermatophagoides farinae  
US-10-218-743-35

Query Match 28.4%; Score 711.5; DB 9; Length 509;  
Best Local Similarity 36.4%; Pred. No. 6.1e-52;  
Matches 148; Conservative 74; Mismatches 152; Indels 33; Gaps 10;

QY 2 QLTCTTWAQYRPGLRMPNDIDPCLCTHLIYAFAGQNNETIENW----- 51  
Db 34 RIVCYVGTWSVYHK-VDPYTTIEDIDPFKCTHLMYGFADIDYKTYIQVDFPQDDNNHNSW 92

QY 52 DVTLYQAFNGLKNKNSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVKFLRQYEF 111  
Db 93 EKHGVERFNNLRKNPELTMTISLGGWEGSEKYSDMAANPTYRQVQSVDLDFLOEYKF 152

QY 112 DGLDFDWEYPGSR-GSPPODKHLFTVLVOEMREAFQEAQINKRPLMVTAAVAGISNI 170  
Db 153 DGLDLDWEYPGSRGKPKIDKQNYLTTLVRELKEAFEPFG-----YLLTAASVPGKDKI 205

QY 171 QSGYEIPOLSOYLDYIHVMTYDLHGSWEGYTGENSPLYKYPTDGTGS-NAYLNVDVYMNW 229  
Db 206 DVAYELKELNQLFDWMNVMTYDHGWNVFVGHNAPLYKRPDETDELHTYFNVTMTHY 265

QY 230 KONGAPAEKLVIGFPTYGHNFIILSNPSTGIGAPTSAGPAGPYAKESGIWAYEIC-TF 288  
Db 266 LNNGATRDKLNVGVPFYGYRAWSIEDRSKVKGDPKAGMSPPGFITGEGVLSYIELCQLF 325

QY 289 LKNGATQGWADAPQEVYAYQGNVWVGYNDIKSFQDKAQLKHNKFGGAMVAIDDDFTG 348  
Db 326 QKEEWHIQDEYNNAPYGYNDIKWGYDDLASISCKLAFKELGVSVMWISLENDDFKG 385

QY 349 TFCNKGKPELSTLKKALG-----LQASCTAPAPQIEPITAA 386  
Db 386 -HCGP-KYPLLKNVHNMINGDEKNSYECLLGPSTTPT-PTTPTSTS 429

## RESULT 10

US-10-218-743-38  
; Sequence 38, Application US/10218743  
; Publication No. US20030096779A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine A.  
; APPLICANT: Hunter, Shirley Wu  
; APPLICANT: Weber, Eric R.  
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
; FILE REFERENCE: AL-2-C3  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR FILING DATE: 1999-04-15  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR FILING DATE: 1998-04-17  
; PRIOR FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Dermatophagoides farinae  
US-10-218-743-38

Query Match 28.4%; Score 711.5; DB 9; Length 509;  
Best Local Similarity 36.4%; Pred. No. 6.1e-52;  
Matches 148; Conservative 74; Mismatches 152; Indels 33; Gaps 10;

QY 2 QLTCTTWAQYRPGLRMPNDIDPCLCTHLIYAFAGQNNETIENW----- 51  
Db 34 RIVCYVGTWSVYHK-VDPYTTIEDIDPFKCTHLMYGFADIDYKTYIQVDFPQDDNNHNSW 92

QY 52 DVTLYQAFNGLKNKNSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVKFLRQYEF 111  
Db 93 EKHGVERFNNLRKNPELTMTISLGGWEGSEKYSDMAANPTYRQVQSVDLDFLOEYKF 152

QY 112 DGLDFDWEYPGSR-GSPPODKHLFTVLVOEMREAFQEAQINKRPLMVTAAVAGISNI 170  
Db 153 DGLDLDWEYPGSRGKPKIDKQNYLTTLVRELKEAFEPFG-----YLLTAASVPGKDKI 205

QY 171 QSGYEIPOLSOYLDYIHVMTYDLHGSWEGYTGENSPLYKYPTDGTGS-NAYLNVDVYMNW 229  
Db 206 DVAYELKELNQLFDWMNVMTYDHGWNVFVGHNAPLYKRPDETDELHTYFNVTMTHY 265



APPLICANT: McCall, Catherine A.  
APPLICANT: Hunter, Shirley Wu  
APPLICANT: Weber, Eric R.  
TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
FILE REFERENCE: AL-2-C3  
CURRENT APPLICATION NUMBER: US/10/218,743  
PRIOR FILING DATE: 2002-08-13  
PRIOR APPLICATION NUMBER: US/09/292,225  
PRIOR FILING DATE: 1999-04-15  
PRIOR APPLICATION NUMBER: 60/098,909  
PRIOR FILING DATE: 1998-09-02  
PRIOR APPLICATION NUMBER: 60/085,295  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/098,565  
PRIOR FILING DATE: 1998-04-17  
PRIOR APPLICATION NUMBER: 09/062,013  
PRIOR FILING DATE: 1998-04-17  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 18  
LENGTH: 555  
TYPE: PRT  
ORGANISM: Dermatophagoides farinae  
US-10-218-743-18

Query Match 28.3%; Score 708.5; DB 9; Length 555;  
Best Local Similarity 35.7%; Pred. No. 1.2e-51;  
Matches 148; Conservative 77; Mismatches 154; Indels 35; Gaps 9;

QY 2 QLTCTFTWAQYRPGGRMPDNIDPCLCTHLIYAFAGQNNETIENW-----51  
DB 34 RIVCYVGTWSVYHK-VDPYTIEDIDPCKTILMYGFAKIDYKTYTQVDPYQDDNNHNSW 92  
QY 52 DVTLYQAFNLGNKNSQLTLLAIGGNFCTAPFTAMVSTPENROFTITSVILKFLRQYEF 111  
DB 93 EKRGRYFNRLKLNKPELTMTISLGGWYEGSEKISDMAANPTTQOFTOSVLDFFLOEYKF 152  
QY 112 DGLDFWMEYFSGR-GSPPODKHLFTVLVQEMREAFQEAQKINKPRLMTAAVAAGISNI 170  
DB 153 DGLDLDWEYFSGRLGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAASPGKDKI 205  
QY 171 QSGYEIPQLSOYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGS-NAYLVNDYVMYTW 229  
DB 206 DRAYDIELKELNKLFDWNVMTYDYGWENFYGHNAPLYRPOEDTDLHTYFNVNMTHTY 265  
QY 230 KONGAPAEKLIYGFPTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEIC-TF 288  
DB 266 LNNATRDVLVGMVFPYGRAWSIEDRSKLGKIDPAKGMSPPGFISGEEGVLSTIELCQLF 325  
QY 289 LKNGATQGWDAFQEVYAYQGNVWGYDNIKSPDIKAQWLKHNKFGGAMVWALDLDFTG 348  
DB 326 QKEEMHIQYDEYNAYGYNDKIWGYDDLASTSCKLAFELKELGVSGVWVWSLENDDFKG 385  
QY 349 TFCNQOKFPLI-----STLKKALGLOSASCTAPAOPIETIAPSGS 390  
DB 386 -HCGP-KNPLLNKVNHWINGDEKNSEPCILGSPSTTTPTTPTTPTTPTTPTTPTT 437

RESULT 14  
US-09-923-844B-2  
Sequence 2, Application US/09923844B  
Patent No. US20020166143A1  
GENERAL INFORMATION:  
APPLICANT: Pioneer Hi-Bred International, Inc.  
APPLICANT: Bao, Zhongmeng  
APPLICANT: Lu, Guihua  
TITLE OF INVENTION: Sclerotinia-inducible Genes and  
TITLE OF INVENTION: Promoters and Their Uses  
FILE REFERENCE: 35718/234631  
CURRENT APPLICATION NUMBER: US/09/923,844B  
CURRENT FILING DATE: 2001-08-07  
PRIOR APPLICATION NUMBER: US 60/224,603

PRIOR FILING DATE: 2000-08-11  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 371  
TYPE: PRT  
ORGANISM: Helianthus annuus  
US-09-923-844B-2

Query Match 18.8%; Score 472; DB 9; Length 371;  
Best Local Similarity 34.1%; Pred. No. 7.5e-32;  
Matches 118; Conservative 47; Mismatches 145; Indels 36; Gaps 12;

QY 6 YETNWAQ-YRPGGRMPDNIDPCLCTHLIYAFAGQNNETIENW-----DVTLYQAFN 60  
DB 30 YWPSWAQDFLP-----PSNIQTAYFTHYVYAFSLSPNN---VTFQFDVHRTTASALNSFN 80  
QY 61 -GLKNKNSQLTLLAIGGNFCTAP-FTAMVSTPENROFTITSVILKFLRQYEFDGLDFPW 118  
DB 81 TALHGNKPNPVTFLSFGGSAGVKQLFSLASSPGSRAAFHSTIQVARNYYFDGADLDM 140  
QY 119 EYPSGRSGPPPODKHLFTVLVQEMREAFQEAQKINKPRLMTAAVAAGISNIQSG---YE 175  
DB 141 EYPTQ-----TDMNFGLLDEWRVAVNNEATSTCKPRLLSAATHEBEVDRDNGVAKYP 196  
QY 176 IPQLSOYLDYIHVMYDYLHGSW-EGYTGNSPLYKYPTDTGSNAYLVNDYVMYWKDNGA 234  
DB 197 VASINKNLGDINAMCYDHYGWPWTPDATGAPAAALY-----NPNGLSTSLNGLSWISAGI 250  
QY 235 PAEKLIYGFPTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTF-LKNGA 293  
DB 251 ORQKLVGMPLYGVTWKLKNPSVNGIGAPAGIGPG-----NEGAMLYSEVOOFNAQNNA 305  
QY 294 TQGWDAPEVPYAYQGNVWGYDNIKSPDIKAQWLKHNKFGGAMVW 339  
DB 306 RVYDTQTVSYISYSGTTWIGYDDVNSVQRKYQYAKSLNIGGYFFW 351

RESULT 15  
US-09-748-033-3  
Sequence 3, Application US/09748033  
Patent No. US20020069431A1  
GENERAL INFORMATION:  
APPLICANT: Broadway, Roxanne M.  
APPLICANT: Gongora, Carmona E.  
TITLE OF INVENTION: EFFECT OF ENDOCHITINASE AND CHITOBIOCIDASE AND THEIR  
TITLE OF INVENTION: ENCODING GENES ON PLANT GROWTH AND DEVELOPMENT  
FILE REFERENCE: 19603/3091  
CURRENT APPLICATION NUMBER: US/09/748,033  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/172,003  
PRIOR FILING DATE: 1999-12-23  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 3  
LENGTH: 376  
TYPE: PRT  
ORGANISM: Streptomyces albidoflavus  
US-09-748-033-3

Query Match 11.4%; Score 285.5; DB 10; Length 376;  
Best Local Similarity 28.0%; Pred. No. 4.7e-16;  
Matches 117; Conservative 43; Mismatches 145; Indels 113; Gaps 25;

QY 6 YETNWAQYRPGGR-FMPDNI-----DPCLCTHLIYAFAGQNNETIETIE-----49  
DB 14 YFTWGVY-----GRNYHVKNLVTSAGKITHINYSFGNVQGGKCTIGDSPFAAYDKAYTA 69  
QY 50 -----WMDVTLYQAFN---GLKNKNSQLTLLAIGGNFCTAPFTAMVSTPENRQ 96  
DB 70 AESVDGVADTW-DQPLRGNFNLRLKAKYPHIKVLVSGGWTW-SGGFTDAVKPAAPA 127  
QY 97 TFIITSVILKFLRQYE-FDGLDFWMEYFSGR-----SPQDKHLFTVLVQEMREAFQEA 149



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:56:54 ; Search time 14.1346 Seconds  
(without alignments)  
3094.613 Million cell updates/sec

Title: US-10-004-219b-14  
Perfect score: 2506  
Sequence: 1 YQLTCYFTNWAQYRPLGRF.....QQNCQAGLVFDTSCDCNWA 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1361.5	54.3	399	2 S27879	secretory protein
2	976.5	39.0	383	2 A49562	cartilage glycopro
3	968.5	38.6	383	2 S51327	heparin-binding gl
4	959.5	38.3	654	2 I38605	oviductal glycopro
5	943.5	37.6	537	2 S57197	oviduct-specific g
6	941.5	37.6	539	2 I46470	estrogen dependent
7	912.5	36.4	483	2 A53918	chitinase (EC 3.2.
8	885	35.3	617	2 T15408	hypothetical prote
9	874	34.9	525	2 T44445	chitinase (EC 3.2.
10	863	34.4	554	2 A56596	chitinase (EC 3.2.
11	801.5	32.0	405	2 S61551	breast-regressing
12	786.5	31.4	504	2 A38221	chitinase (EC 3.2.
13	768.5	30.7	1635	2 T14075	chitinase (EC 3.2.
14	488.5	19.5	1215	2 T43916	chitinase A [impor
15	482	19.2	599	2 D83764	chitinase BH0916 [
16	472.5	18.9	423	2 JQ1975	chitinase (EC 3.2.
17	465.5	18.6	398	2 T04761	chitinase homolog
18	463.5	18.5	424	2 S47133	chitinase (EC 3.2.
19	463.5	18.5	699	2 A38368	chitinase (EC 3.2.
20	457.5	18.3	1484	2 T29275	hypothetical prote
21	456.5	18.2	756	2 AB1452	chitinase B homolo
22	454.5	18.1	2025	2 T03884	hypothetical prote
23	451.5	18.0	756	2 AB1088	chitinase B homolo
24	450	18.0	831	2 T00323	chitinase (EC 3.2.
25	444.5	17.7	379	2 T04762	chitinase homolog
26	440.5	17.6	423	2 S68121	chitinase I precur
27	434.5	17.3	423	2 S51369	chitinase - fungus
28	427.5	17.1	427	2 JC4565	chitinase (EC 3.2.
29	427	17.0	546	2 F84238	chitinase [importe

30	402	16.0	378	2 S51591	chitinase (EC 3.2.
31	393	15.7	366	2 T04763	chitinase homolog
32	388.5	15.5	511	2 S61166	probable membrane
33	388	15.5	452	2 JC4038	47K glycoprotein p
34	374.5	14.9	371	2 T04756	chitinase homolog
35	369.5	14.7	365	2 T04757	chitinase homolog
36	369.5	14.7	1051	2 D82428	chitodextrinase VC
37	368	14.7	633	2 T24898	hypothetical prote
38	367.5	14.7	869	2 T44440	chitinase (EC 3.2.
39	357.5	14.3	849	2 D82510	chitinase VCA0027
40	352.5	14.1	1054	2 T30933	chitinase (EC 3.2.
41	349.5	13.9	563	2 S60651	chitinase precursor
42	347.5	13.9	499	2 S52422	chitinase (EC 3.2.
43	345.5	13.8	332	2 T04754	hypothetical prote
44	344.5	13.7	499	2 S04856	chitinase (EC 3.2.
45	343	13.7	1046	2 T30199	chitinase (EC 3.2.

ALIGNMENTS

RESULT 1

S27879  
secretory protein YM-1 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 22-Jun-1999  
C:Accession: S27879  
R:Chang, N.C.A.; Liu, C.H.; Chang, A.C.  
submitted to the EMBL Data Library, June 1992  
A:Description: Molecular characterization of a secretory protein (YM-1) transiently expre  
A:Reference number: S27879  
A:Accession: S27879  
A:Molecule type: mRNA  
A:Residues: 1-399 <CHA>  
A:Cross-references: EMBL:M94584; NID:G202441; PIDN:AAB62394.1; PID:G202442  
C:Superfamily: Streptomyces chitinase chi40  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-399/Product: secretory protein YM-1 #status predicted <MAT>

Query Match 54.3%; Score 1361.5; DB 2; Length 399;  
Best Local Similarity 66.3%; Pred. No. 7.3e-95;  
Matches 250; Conservative 46; Mismatches 80; Indels 1; Gaps 1;

QY	1	YQLTCYFTNWAQYRPLGRFMPDNIDPCLCTHLIYAFAGRQNNETITIEWNDVLYQAFN	60
DB	22	YQLMCCYTSWAKDRPIEGSFKEGNDPCLCTHLIYAFAGMNNETIYTHEQDLRDYEALN	81
QY	61	GLKN-KNSQLTKLLAIGGNFCTAPFTAMVSTPENRQFTITSVIKFLQYEPDGLDFWE	119
DB	82	GLKDKKNTLTKLLAIGGWKFGPAPFAMVSTPQNRQIFQSVIRFLQYRNPDLNLDWQ	141
QY	120	YPSRSGSPQDKHLFTVLVOEMREAFQEAQINKPRLMTVAAGISNIQSGVEIPQL	179
DB	142	YPSRSGSPQDKHLFTVLVOEMREAFQEAQINKPRLMTVAAGISNIQSGVEIPQL	201
QY	180	SOYLDYHWTYDLHGSGWEGYTGNSPLYKYPTDGTGSNAYLVNDYVMYKNDGAPAEKL	239
DB	202	LSLLDYIQWTVYDLHDPKDYTGNSPLYKSPYDYGKSLADLNVDISIYWKDHGAASEKL	261
QY	240	IVGFPTYGHNFILSNPSNTGICAPTSAGAPAGYAKESGIWAYYEICTPLKNGATQGWDA	299
DB	262	IVGFPAYGHTFILDSPSKTGIGAPTISTGPPGKYTDESGLLAYVEVCTPLNKGATEVMDA	321
QY	300	POEVPYAGVQNVGVYDNIKSPDIKAQMLKHNKFGMVMYALDLDFTCTFCNGQGFPLI	359
DB	322	POEVPYAGVQNVGVYDNIKSPDIKAQMLKHNKFGMVMYALDLDFTCTFCNGQGFPLI	381
QY	360	STLKKALGLQASCTAP	376
DB	382	STLKGDLNHSASCKGP	398

RESULT 2

A49562  
cartilage glycoprotein gp39 precursor - human  
N;Alternate names: 39K synovial protein  
C;Species: Homo sapiens (man)  
C;Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 22-Jun-1999  
C;Accession: A49562; #sequence\_revision 23-Mar-1995 #text\_change 22-Jun-1999  
R;Hakala, B.E.; White, C.; Recklies, A.D.  
J. Biol. Chem. 268, 25803-25810, 1993  
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes and  
A;Reference number: A49562; MUID:94064658; PMID:8245017  
A;Accession: A49562  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-383 <HA>  
A;Cross-references: GB:M80927; NID:G348911; PIDN:AAA16074.1; PID:G348912  
R;Nyikos, P.; Golds, E.E.  
Biochem. J. 269, 263-268, 1990  
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary prote  
A;Reference number: S10677; MUID:90328983; PMID:2375755  
A;Accession: S10677  
A;Molecule type: protein  
A;Residues: 22-40, 'X', 42-45 <NY2>  
C;Superfamily: Streptomyces chitinase chi40  
C;Keywords: cartilage; extracellular protein; glycoprotein  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>  
  
Query Match 39.0%; Score 976.5; DB 2; Length 383;  
Best Local Similarity 51.2%; Pred. No. 6.7e-66;  
Matches 188; Conservative 62; Mismatches 108; Indels 9; Gaps 5;  
  
QY 1 YOLTCTFTNWAQYRPGCLGRFMPDNIDPCLCTHLIYAFAGRQNNETITIEWNDVTLQAFN 60  
DB 22 YKLVCTYTSWSQYREGDSCFPDLPFLCTHLIIYFANISNNEIDTLEWNDVTLQAFN 81  
  
QY 61 GLKNKNSQLKTLAIGGNWFGTAPPTAMVSTPENRQTFTSVIKFLQYEFGLDFDWEY 120  
DB 82 TLKRNPNLKTLLSVGGWNGFSQRFKASNTQSRRTIKSVPPFLRTHGFDGLDAMI 141  
  
QY 121 PGRSGPPQDKHLFTVLVQEMREAFQEAQKQINPRLMTAAVAAGISNIQSGYIPQLS 180  
DB 142 PGRS-----DKQHFTTLIKEMKAEFTKEA-QPGKKQLLSAALSAGKVTIDSSYDI 195  
  
QY 181 QYLDYIHVMYDHLGSGWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKNGAPAEKLI 240  
DB 196 QHLDFTISIMTYDFHGAWRGTTGHSPLFRQGDASDRFSNADYAVSVLRGLGAPASKLV 255  
  
QY 241 VGPPTVGHNFILSNPTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300  
DB 256 MGIPTFGRSFTLAS-SETGVGAPISGPGIPGRFTKEAGTLAYEICDFLR-GATVHR 313  
  
QY 301 QVPYAYQGNVWGYDNIKSFDIKAWLKHNGKGGAMWALDLDFTGTCNQG-KFPLI 359  
DB 314 QQVPYATKGNQWGYDDQESVKSQVYLKDRQLAGAMWALDLDFTGSGFCQDURFPLT 373  
  
QY 360 STLKAL 366  
DB 374 NAIKDAL 380  
  
RESULT 3  
S51327  
heparin-binding glycoprotein 38K - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 22-Jun-1999  
C;Accession: S51327  
R;Shackleton, L.M.; Mann, D.M.; Millis, A.J.T.  
submitted to the EMBL Data Library, January 1995  
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differe  
A;Reference number: S51327  
A;Accession: S51327  
A;Status: preliminary  
A;Molecule type: mRNA

A;Residues: 1-383 <SHA>  
A;Cross-references: EMBL:247803; NID:G634097; PIDN:CAA87764.1; PID:G634098  
C;Superfamily: Streptomyces chitinase chi40  
  
Query Match 38.6%; Score 968.5; DB 2; Length 383;  
Best Local Similarity 51.5%; Pred. No. 2.7e-65;  
Matches 189; Conservative 57; Mismatches 112; Indels 9; Gaps 5;  
  
QY 1 YOLTCTFTNWAQYRPGCLGRFMPDNIDPCLCTHLIYAFAGRQNNETITIEWNDVTLQAFN 60  
DB 22 YKLVCTYTSWSQYREGDSCFPDLPFLCTHLIIYFANISNNEIDTLEWNDVTLQAFN 81  
  
QY 61 GLKNKNSQLKTLAIGGNWFGTAPPTAMVSTPENRQTFTSVIKFLQYEFGLDFDWEY 120  
DB 82 TLKRNPNLKTLLSVGGWNGFSQRFKASNTQSRRTIKSVPPFLRTHGFDGLDAMI 141  
  
QY 121 PGRSGPPQDKHLFTVLVQEMREAFQEAQKQINPRLMTAAVAAGISNIQSGYIPQLS 180  
DB 142 PGRS-----DKRHFTTLVKEMKAEFVREALP-GTERLLLSGAVSAGKVAIDRGYDIAQIS 195  
  
QY 181 QYLDYIHVMYDHLGSGWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKNGAPAEKLI 240  
DB 196 QHLDFTISIMTYDFHGAWRGTTGHSPLFRQGDASDRFSNADYAVSVLRGLGAPASKLV 255  
  
QY 241 VGPPTVGHNFILSNPTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300  
DB 256 MGIPTFGRSFTLAS-SKTDVGAPASGPGIPGRFTKEKGLAYEICDFLQ-GATVRRPLG 313  
  
QY 301 QVPYAYQGNVWGYDNIKSFDIKAWLKHNGKGGAMWALDLDFTGTCNQG-KFPLI 359  
DB 314 QQVPYATKGNQWGYDDQESVKNKAKYLSKRLAGAMWALDLDFTGFGNFCQNLRFPLT 373  
  
QY 360 STLKAL 366  
DB 374 SAIKDL 380  
  
RESULT 4  
I38605  
oviductal glycoprotein - human  
C;Species: Homo sapiens (man)  
C;Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 29-May-1998  
C;Accession: I38605  
R;Arias, E.B.; Verhage, H.G.; Jaffe, R.C.  
Biol. Reprod. 51, 685-694, 1994  
A;Title: Complementary deoxyribonucleic acid cloning and molecular characterization of a  
A;Reference number: I38605; MUID:95119256; PMID:7819450  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-654 <RES>  
A;Cross-references: EMBL:U09550; NID:G529147; PID:G529148  
  
Query Match 38.3%; Score 959.5; DB 2; Length 654;  
Best Local Similarity 49.3%; Pred. No. 2.6e-64;  
Matches 187; Conservative 56; Mismatches 129; Indels 7; Gaps 3;  
  
QY 1 YOLTCTFTNWAQYRPGCLGRFMPDNIDPCLCTHLIYAFAGRQNNETITIEWND-VTLQAF 59  
DB 22 HKLVCTYTNWAHSRFGPASILPHDLPFLCTHLIFAFASMNNOQIVAKDQEKILYPEF 81  
  
QY 60 NGLKNKNSQLKTLAIGGNWFGTAPPTAMVSTPENRQTFTSVIKFLQYEFGLDFDWEY 119  
DB 82 NKLKRNPNLKTLLSVGGWNGFSQRFKASNTQSRRTIKSVPPFLRTHGFDGLDAMI 141  
  
QY 120 YPGRSGPPQDKHLFTVLVQEMREAFQEAQKQINPRLMTAAVAAGISNIQSGYIPQL 179  
DB 142 YPGLRSGPMHWRWTFLFLIEELLFAFRKEALLTMRPLLSSAASGVPHIVQTSYDVRFL 201  
  
QY 180 SQYLDYIHVMYDHLGSGWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKNGAPAEKLI 239  
DB 202 GRLLDFINVLSDYDLHGSWERFTGHNSPLFSPLFDPKSSA-----YAMNYWRKLGAPSEKL 256



QY 240 IVGFPTYGHNFTLSNPSNTGIGAPTSAGPAGYAKESGIWAYYEICTFLKNGATQGWDA 299  
Db 257 IMGITYGRTFRLKAKSNGLOARIGASPCKYTKQSGFLAYFEICSPVW-GAKKHWD 315  
QY 300 POEVPYAVGNNVWVGYDNKSPDIKAQMLKHNKFGAMVWALDLDFTGTCNQGKFPPLI 359  
Db 316 YQYVFPYANKGEWGYDNAISFYKAWFIRRHFGAMVWTLMDMDVDRGTGTCGTPPPLV 375  
QY 360 STLKKALGLQASCTAPAQ 378  
Db 376 YVNDILVRAEFSSTSLPQ 394  
RESULT 5  
S57197  
oviduct-specific glycoprotein 95K precursor - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 05-Nov-1999  
C:Accession: S57197  
R:Sendai, Y.; Abe, H.; Kikuchi, M.; Satoh, T.; Hoshi, H.  
Biol. Reprod. 50, 927-934, 1994  
A:Title: Purification and molecular cloning of bovine oviduct-specific glycoprotein.  
A:Reference number: S57197; MUID:94257768; PMID:8199272  
A:Accession: S57197  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-537 <SEN>  
A:Cross-references: EMBL:D16639; NID:g391621; PID:BAA04065.1; PID:d1004583; PID:g391622  
C:Keywords: glycoprotein  
Query Match 37.6%; Score 943.5; DB 2; Length 537;  
Best Local Similarity 49.6%; Pred. No. 3.2e-63;  
Matches 182; Conservative 53; Mismatches 125; Indels 7; Gaps 3;  
QY 1 YOLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTEWND-VTLYQAF 59  
Db 19 HKLVCFYTNWAFSRPGASILPRDLDPFLCTHLVFAFASMSNNQIVPKDPQDEKILYPEF 78  
QY 60 NGLKKNKSQLTKLLAIGWNFGCTAPFTAMVSTPENROTFTITSVIKFLROYEDGLDFWE 119  
Db 79 NKLKERNRGLTKLLSIGWNFGCTVFTMTLSFRERFVSSVIALRLTHGDFGLDLFFL 138  
QY 120 YPGSRGSPQDKHLFTVLVQEMREAFEOEAKOINKPRLMVTAAVAGISNIQSGYEIPQL 179  
Db 139 YPLGSGPARDRWTFVLELLQAFKNEAQLTWPRLLLSAASVGDHPHVQKAYEARLL 198  
QY 180 SOYLDYIHVMYTDLHGSWEYTGNSPLYKYPTDTGNSNAYLVNVDYVMYWKDNGAPAEKL 239  
Db 199 GRLLDFISVLSYDLHGSWEKVTGHSNPLSLPGDPKSSA-----YAMSYWRQLGVPPPEKL 253  
QY 240 IVGFPTYGHNFTLSNPSNTGIGAPTSAGPAGYAKESGIWAYYEICTFLKNGATQGWDA 299  
Db 254 LMGLPTYGRTFHLKASQNELRAQAVGPASPKYTKQAGFLAYEICEFVR- AKKRWIN 312  
QY 300 POEVPYAVGNNVWVGYDNKSPDIKAQMLKHNKFGAMVWALDLDFTGTCNQGKFPPLI 359  
Db 313 DQYVPYAFKGEWGYDDAISFGYKAFKIKREHFGAMVWTLDDDFRGYFCGTGPPPLV 372  
QY 360 STLKKAL 366  
Db 373 HTLNLL 379  
RESULT 6  
I46470  
estrogen dependent oviduct protein precursor - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I46470  
R:Desouza, M.M.; Murray, M.K.  
Endocrinology 136, 2485-2496, 1995  
A:Title: An estrogen-dependent secretory protein, which shares identity with chitinases,  
and embryo development.

A:Reference number: I46470; MUID:95269691; PMID:7750470  
A:Accession: I46470  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-539 <DES>  
A:Cross-references: EMBL:U16719; NID:9885600; PIDN:AAC48471.1; PID:g885601  
Query Match 37.6%; Score 941.5; DB 2; Length 539;  
Best Local Similarity 48.8%; Pred. No. 4.6e-63;  
Matches 179; Conservative 59; Mismatches 122; Indels 7; Gaps 3;  
QY 1 YOLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTE-WNDVTLYQAF 59  
Db 22 HKLVCFYTNWAFSRPGASILPRDLDPFLCTHLVFAFASMSNNQIVPKDPLEKILYPEF 81  
QY 60 NGLKKNKSQLTKLLAIGWNFGCTAPFTAMVSTPENROTFTITSVIKFLROYEDGLDFWE 119  
Db 82 NKLKERNRGLTKLLSIGWNFGTSRFTKMLSTFNSRERFVSVIALRLTHGDFGLDLFFL 141  
QY 120 YPGSRGSPQDKHLFTVLVQEMREAFEOEAKOINKPRLMVTAAVAGISNIQSGYEIPQL 179  
Db 142 YPLGSGPARDRWTFVLELLQAFKNEAQLTWPRLLLSAASVGDHPHVQKAYEARLL 201  
QY 180 SOYLDYIHVMYTDLHGSWEYTGNSPLYKYPTDTGNSNAYLVNVDYVMYWKDNGAPAEKL 239  
Db 202 GRLLDFISVLSYDLHGSWEKVTGHSNPLSLPGDPKSSA-----YAMSYWRQLGVPPPEKL 256  
QY 240 IVGFPTYGHNFTLSNPSNTGIGAPTSAGPAGYAKESGIWAYYEICTFLKNGATQGWDA 299  
Db 257 LMGLPTYGRTFHLKASQNELGAAGSPASPKYTKQAGFLAYEYEVCSFVOR-AKKRWIN 315  
QY 300 POEVPYAVGNNVWVGYDNKSPDIKAQMLKHNKFGAMVWALDLDFTGTCNQGKFPPLI 359  
Db 316 DQYVPYAFKGEWGYDDAISFGYKAFKIKREHFGAMVWTLDDDFRGNGCGTGPFPPLA 375  
QY 360 STLKKAL 366  
Db 376 HTLNLL 382  
RESULT 7  
A53918  
chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)  
C:Species: Chelonus sp.  
C:Date: 28-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 21-Jul-2000  
C:Accession: A53918  
R:Krishnan, A.; Nair, P.N.; Jones, D.  
J. Biol. Chem. 269, 20971-20976, 1994  
A:Title: Isolation, cloning, and characterization of new chitinase stored in active form  
A:Reference number: A53918; MUID:94342256; PMID:8063715  
A:Accession: A53918  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-483 <KRI>  
A:Cross-references: GB:U10422; NID:g533504; PIDN:AAA61639.1; PID:g533505  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
Query Match 36.4%; Score 912.5; DB 2; Length 483;  
Best Local Similarity 38.0%; Pred. No. 6.1e-61;  
Matches 185; Conservative 82; Mismatches 157; Indels 63; Gaps 11;  
QY 2 QLTVCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTE-WND---VTLY 56  
Db 23 KVCYFGAWSVTRQNGKFDINGIDFTLCTHLIYSGVGNCKVQVLDPSDLPNLDGF 82  
QY 57 QAFNGLKKNKSQLTKLLAIGWNFGCTAPFTAMVSTPENROTFTITSVIKFLROYEDGLDF 116  
Db 83 GKFTSLRKNKNSVKIMVAVGWNAGSVFSSQWASDQATREAFQNVKFLQYQYQDFGDI 142  
QY 117 DWYPGSRGSPQDKHLFTVLVQEMREAFEOEAKOINKPRLMVTAAVAGISNIQSGYEI 176  
Db 143 DWYEPQARGGSPADYKNNVMVGLCKALKKAFVQH-----DYILSAVAAPETSASKSYDI 195

```

Db      402  NAGCSNSNGLYPLISVIAKELGSGVIIPKGGVVTAPTATTATTGTP--PMTSAVTTT 459
Qy      391  GNGSGSSGG---SSGSGGFCVAVRANGLYPVANNRFAFVHCVNGVTYQQNCQAGLVFDT 447
Db      460  TAATTITTRAATTATNTNVCSGKSDGFYFNSNNCGFLVLCSSKSYSMSCPSGLQYSA 519
Qy      448  SCDC 452
Db      520  SLKYC 524

RESULT 9
T4445
chitinase (EC 3.2.1.14) [imported] - African malaria mosquito
C;Species: Anopheles gambiae (African malaria mosquito)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44445
R;Shen, Z.; Jacobs-Lorena, M.
submitted to the EMBL Data Library, June 1997
A;Reference number: Z22771
A;Accession: T44445
A;Status: preliminary; translated from GE/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-525 <SHE>
A;Cross-references: EMBL:AF008575; PIDN:AAB87764.1
A;Experimental source: adult; gut
C;Genetics:
A;Gene: chi-1
C;Keywords: glycosidase; hydrolase

Query Match      34.98; Score 874; DB 2; Length 525;
Best Local Similarity 37.38; Pred. No. 5.4e-58;
Matches 188; Conservative 84; Mismatches 164; Indels 68; Gaps 14

Qy      2  QLTCTFNWAQYRPGLRFPMDIPCLCTHLIYAFAGRONNETTIENNDVTL----- 55
Db      32  KVCYVGTWAVYRFGNGRYDIEHDIPSLCTHLMYGFPG--INEDATVRIIDPYLDLENN 89
Qy      56  ----YQAFNGLKKNNSOLKTLAIGGWNFGTAPPTAMVSTPENRQTITSVIKELRQYEF 111
Db      90  GRGHIRFVGLKKNVGLKLTAAIAGGWNESGRFSAAMAASGELKRFISDCVAFQCRHGF 149

```

QY	112	DGJDFDWEYFGSR-GSPFPQDKHLFTVLVQEMRKAPEQEAQINAPRLMWTAAVA-----A	166
Db	150	DGIDLWEYPAQDGNPLIDRDNHAQLVEBMRBEFDHYG-----LLLTAASVAFPSA	202
QY	166	GISNIOGSEIPIOLSOQLDYIHWMTYDLHGSWEGYTGENSEPLYKYPTDTSN-AVLNVDY	224
Db	203	GVS-----YDIPRISKSFHFLVMVYDMHGAWDSYCGINAPLYRGSDATTDRLQGINVA	257
QY	225	VMMYWKDNGAPAKLIVFPPTYGHNFTLSPNSNTGICAPTSGAGCAPGAKESGIWAYYE	284
Db	258	SIHFWLAQOQGTGRKLVLGIPLYGRNFTLASAANTQIGAPTVGCGTVGRYTRPVGVMGYNE	317
QY	285	ICTFLKNGATQOWD----APQEVYPYAVQGNVMVGYDNIKSFDIKAQWLKHKNFCGAMVWA	340
Db	318	FCBKL---ATEAWDLRWSEQQQPYAVRNQWGYDDLRSSVQLKVKYLLDQGLGGAMVWS	374
QY	341	IDLDDEFTGFCNOGKEPLISTLAKALGLQASCTAPAQIETITA-----APSGSG	391
Db	375	LETDDFLGV-CGGGRYPLMHEIRSLVNGGTPTSTTTPPSPVAPTSTVAPGTTTTTPTGAN	433
QY	392	NG-----SGSSSSGGSSG-----SGGFCVAVRANGLYPVANNRNNAFHWCVNG	432
Db	434	PGTQPTSDAPHNHTTSTTEGPNPQTRTPPSGDGPCAGGRYGVFPHTNCARYIICLTA	493
QY	433	VTYQQ-NCQAGLVFDTSCDCNWA	455
Db	494	DTYYETCPGTLFDPALHICNWA	517

RESULT 10





C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: D83764  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maqui, N.; Fuji, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: D83764  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-599 <STO>  
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04635.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH916

	Query Match	19.2%;	Score 482;	DB 2;	Length 599;
	Best Local Similarity	27.9%;	Pred. No. 2.1e-28;		
	Matches 127;	Conservative 85;	Mismatches 144;	Indels 100;	Gaps 20;
Qy	1	YQLTCYTNWAOYRPLGR-FMPDNIDPCLCTHLIYAF-----GRQN-----	43		
Db	40	YKIVAYPSWGAY-----GRDQVWDIDASKISHINYAFANICWDGRHGNPDGPNQW 95			
Qy	44	-----BITTIE-----WNVTL-----YQAFNGLKNKNSQLKTLA 74			
Db	96	SCQDENGVIDVPNGSIVMGDPWIDAKSNPGDTWDEPLRGNFQKLNKLEHPLKTLIS 155			
Qy	75	IGGNWFGTAPTAMVSTPENRQTETITSVKELROYEFDGLDDEWYVGRGSP-----PQ 129			
Db	156	VGGWTWSNR-FSDMAATKETRENFANSVEFIRKYGFDGVDVDWEYVPSGGLPGNSRPE 214			
Qy	130	DKHLFTVLVQEMREAFEOEAKQINKPRLMVTAAVAAGISNIQSGY-----EIPQLSQYLDY 185			
Db	215	DKENHVLLQEVRLKDEAGQEDGKDLL---TIASGAS---PGIVENKLNIEIAEIVDW 268			
Qy	186	IHMVTDYDLHGSWEGVTGENSPLYKYP-----TDTGSNAVLYVDYVNVYKNDGAPAEKLIV 241			
Db	269	INIMTYDFNGGQWNIASHNAPLYDPATANTELPTEHFNVESAVEGHLQAGVPEHKLVL 328			
Qy	242	GFPTYGHNFNPSNTG-----IGAPTSAGPAGPYAKESGIWAYYEICTFL--KNGATQ 295			
Db	329	GMPFYGRGWSNCDGANQGEYORCAPREGTW-----ENGVFDSLDHYYINKNGYOR 381			
Qy	296	GWDAPQEVPIYAY---QGNVWGVYDNIKFDIKAOWLKKNKFGGAMWALDLDFTGTFCN 352			
Db	382	YNDVAKVPFLYNATNGN-FITYDDEESFRYKTDYFKSNLAGSNFWDVS-GDRNGT--- 436			
Qy	353	QCKFPLISTLKKALQASCTAPAPQIEPIAAPS 388			
Db	437	-----LLTALADQLGFTPEHQEPEEP-----SSAPT 463			

Search completed: June 29, 2003, 21:02:42  
Job time : 15.1346 secs

**THIS PAGE BLANK (0870)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: June 29, 2003, 20:55:24 ; Search time 25.5769 Seconds  
(without alignments)  
3665.470 Million cell updates/sec

Title: US-10-004-219b-14  
Perfect score: 2506  
Sequence: 1 YQLTCVFTNWAQYRGLGRF.....QQNCOAGLVFTSDCCNWA 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.21.\*  
1: sp archea.\*  
2: sp bacteria.\*  
3: sp fungi.\*  
4: sp human.\*  
5: sp invertebrate.\*  
6: sp mammal.\*  
7: sp mhc.\*  
8: sp organelle.\*  
9: sp phase.\*  
10: sp plant.\*  
11: sp rodent.\*  
12: sp virus.\*  
13: sp vertebrate.\*  
14: sp unclassified.\*  
15: sp virus.\*  
16: sp bacteriaph.\*  
17: sp archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	2506	100.0	476 4	Q9BZP6
2	2097.5	83.7	472 11	Q9JLN1
3	2097.5	83.7	473 11	Q9D803
4	2089.5	83.4	473 11	Q9SPH2
5	2039	81.4	472 6	Q9SM17
6	2010	80.2	368 4	Q9ULY4
7	1725	68.8	315 4	Q9ULY3
8	1670.5	66.7	365 11	Q9IXA9
9	1415	56.5	398 11	Q35744
10	1407	56.1	398 11	P70201
11	1387	55.3	398 11	Q9LZ98
12	1384	55.2	402 11	Q8VHG1
13	1379	55.0	402 11	Q8VH43
14	1340.5	53.5	488 13	Q9W334
15	1323	52.8	466 4	Q13231
16	1239	49.4	262 11	Q9D7W6

17	1173	46.8	387 4	Q9H3V8
18	1008.5	40.2	696 5	Q9VZV2
19	980	39.1	1013 5	Q960M0
20	974.5	38.9	383 4	Q96H17
21	972.5	38.8	383 6	Q8SPQ0
22	968.5	38.6	383 6	Q29411
23	966.5	38.6	390 4	Q96F97
24	966	38.5	260 11	Q61201
25	954.5	38.1	624 6	Q19118
26	948.5	37.8	352 11	Q9WTV1
27	933	37.2	457 6	Q95LB3
28	919.5	36.7	396 11	Q9D7Q1
29	912.5	36.4	483 5	Q23737
30	890.5	35.5	4498 5	Q9W223
31	881	35.2	332 6	Q18949
32	879.5	35.1	381 11	Q99J84
33	874	34.9	525 5	Q44079
34	861	34.4	462 5	Q8SYH0
35	855	34.1	544 5	Q9GQC4
36	853	34.0	460 5	Q9W2M7
37	853	34.0	566 5	Q8WR52
38	852	34.0	543 5	Q8GV05
39	852	34.0	565 5	P90710
40	851	34.0	543 5	Q9GR93
41	849.5	33.9	553 5	P91731
42	848	33.8	552 5	Q9GV44
43	832	33.2	565 5	Q9PG9
44	828	33.0	520 5	Q17100
45	822	32.8	524 5	Q27919

## ALIGNMENTS

RESULT 1  
Q9BZP6  
ID Q9BZP6 PRELIMINARY; PRT; 476 AA.  
AC Q9BZP6;  
DC 01-JUN-2001 (TREMREL. 17, Created)  
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)  
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)  
DE Acidic mammalian chitinase precursor (EC 3.2.1.14).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=STOMACH, AND LUNG;  
RX MEDLINE=21125893; PubMed=11085997;  
RA Boot R.G., Blommaert E.F.C., Swart E., Ghauharali-Van Der Vlugt K.,  
RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;  
RT Identification of a Novel Acidic Mammalian Chitinase Distinct from  
RT Chitotriosidase.";  
RL J. Biol. Chem. 276:6770-6778 (2001).  
DR EMBL; AF290004; AAG60019.1;  
DR InterPro; IPR001579; Chitinase 18/2.  
DR InterPro; IPR002557; Chitin bind PerA.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF01607; CBM\_14; 1.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR SMART; SM00494; ChtBD2; 1.  
DR PROSITE; PS01095; CHITINASE\_18; UNKNOWN\_1.  
KW Glycosidase; Hydrolase; Signal.  
FT SIGNAL  
SQ SEQUENCE 476 AA; 52271 MW; 92B27BAD2FEB4CC CRC64;

Query Match 100.0%; Score 2506; DB 4; Length 476;  
Best Local Similarity 100.0%; Pred. No. 1.5e-171;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 YQLTCVFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRQNNETTIEWNDVTLYQAFN 60





```

DR MGD; MGI:1932052; Chitinase_18/2.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 473 AA; 52003 MW; 333874477476695 CRC64;

Query Match      83.7%; Score 2097.5; DB 11; Length 473;
Best Local Similarity 81.5%; Pred. No. 2.8e-142;
Matches 370; Conservative 40; Mismatches 41; Indels 3; Gaps 1;

QY 1 YQLTCYFTNWAQYRGLGRFMDNIDPCLCTHLIYAFAGQNNNEITTIENWNTVLYQAFN 60
DB 22 YNLICYFTNWAQYRGLGSKFPDDINPCLCTHLIYAFAGQNNNEITTIENWNTVLYKAFN 81
QY 61 GLKKNNSQLTKLLAIGGWNFGTAPFTAMVSTPENRQTFTITSVIKELQVYFGLDLDWEY 120
DB 82 DLKRNNSKLTKLLAIGGWNFGTAPFTTWTSTQNRQTFTITSVIKELQVYFGLDLDWEY 141
QY 121 PGRSGSPQDKHLFTVLVOEMREAFQEAQKINPRLMVTAAVAGISNIQSGYIPLQS 180
DB 142 PGRSGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGISNIQAGYIPELS 201
QY 181 QYLDVIHVTYDHLGSGWEGYTGENSEPLYKYPTDGTGNSNAYLVNDYVWYNNKNGAPAEKLI 240
DB 202 KYLDVIHVTYDHLGSGWEGYTGENSEPLYKYPTETGNSNAYLVNDYVWYNNKNGAPAEKLI 261
QY 241 VGFPYGHNFILSPSNTGIGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATQGWADP 300
DB 262 VGFPYGHNFILSPSNTGIGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATQGWADP 321
QY 301 QEVPIYAGQNVWVGYDNIKSFDIKAQWLKHKNKFGGAMVAIDLDDPTGTFNCGKPPPLIS 360
DB 322 QEVPIYAYKANWLGVDNIKSFVSKAQLKQNNFGGAMIWAIDLDDPTGTFNCGKPPPLIS 381
QY 361 TLKKAIGLOSASCTAPAOPIETAPSGSGNGSGSSSGSGSGSGGFCACVANGLYPVA 420
DB 382 TLNKAIGLSTEGCTAPDVPSEPVTPPP---GSGSGSGSGSGSGSGSGGFCADKADGLYPVA 438
QY 421 NNRNAFWHCVNGVTYQONCOAGLVFDTSCDCNCW 454
DB 439 DNRNFAWQCINGITTYQHCQAGLVFDTSCNCCNW 472

RESULT 4
Q99PH2 ID Q99PH2 PRELIMINARY; PRT; 473 AA.
AC Q99PH2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Acidic mammalian chitinase (EC 3.2.1.14).
GN CHIA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=LUNG;
RX MEDLINE=21125893; PubMed=11085997;
RA Boot R.G., Blommaert E.F.C., Swart E., Ghauharali-Van Der Vlugt K.,
RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;
RT Identification of a Novel Acidic Mammalian Chitinase Distinct from
RT Chitotriosidase.
RL J. Biol. Chem. 276:6770-6778(2001).
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
HYDROLASES).

DR MGD; MGI:1932052; Chitinase_18/2.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 473 AA; 51977 MW; 389D87557BEC0784 CRC64;

Query Match      83.4%; Score 2089.5; DB 11; Length 473;
Best Local Similarity 81.3%; Pred. No. 1.1e-141;
Matches 369; Conservative 40; Mismatches 42; Indels 3; Gaps 1;

QY 1 YQLTCYFTNWAQYRGLGRFMDNIDPCLCTHLIYAFAGQNNNEITTIENWNTVLYQAFN 60
DB 22 YNLICYFTNWAQYRGLGSKFPDDINPCLCTHLIYAFAGQNNNEITTIENWNTVLYKAFN 81
QY 61 GLKKNNSQLTKLLAIGGWNFGTAPFTAMVSTPENRQTFTITSVIKELQVYFGLDLDWEY 120
DB 82 DLKRNNSKLTKLLAIGGWNFGTAPFTTWTSTQNRQTFTITSVIKELQVYFGLDLDWEY 141
QY 121 PGRSGSPQDKHLFTVLVOEMREAFQEAQKINPRLMVTAAVAGISNIQSGYIPLQS 180
DB 142 PGRSGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGISNIQAGYIPELS 201
QY 181 QYLDVIHVTYDHLGSGWEGYTGENSEPLYKYPTDGTGNSNAYLVNDYVWYNNKNGAPAEKLI 240
DB 202 KYLDVIHVTYDHLGSGWEGYTGENSEPLYKYPTETGNSNAYLVNDYVWYNNKNGAPAEKLI 261
QY 241 VGFPYGHNFILSPSNTGIGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATQGWADP 300
DB 262 VGFPYGHNFILSPSNTGIGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATQGWADP 321
QY 301 QEVPIYAGQNVWVGYDNIKSFDIKAQWLKHKNKFGGAMVAIDLDDPTGTFNCGKPPPLIS 360
DB 322 QEVPIYAYKANWLGVDNIKSFVSKAQLKQNNFGGAMIWAIDLDDPTGTFNCGKPPPLIS 381
QY 361 TLKKAIGLOSASCTAPAOPIETAPSGSGNGSGSSSGSGSGSGGFCACVANGLYPVA 420
DB 382 TLNKAIGLSTEGCTAPDVPSEPVTPPP---GSGSGSGSGSGSGSGSGGFCADKADGLYPVA 438
QY 421 NNRNAFWHCVNGVTYQONCOAGLVFDTSCDCNCW 454
DB 439 DNRNFAWQCINGITTYQHCQAGLVFDTSCNCCNW 472

RESULT 5
Q95M17 ID Q95M17 PRELIMINARY; PRT; 472 AA.
AC Q95M17;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chitin binding protein b04.
GN CBP B04.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=21475601; PubMed=11591385;
RA Suzuki M., Morimatsu M., Yamashita T., Iwanaga T., Syuto B.;
RT "A novel serum chitinase that is expressed in bovine liver.";
RL FEBS Lett. 506:127-130(2001).
DR EMBL; AB051629; BAB71805.1; -.
DR InterPro; IPR001579; Chitinase_18/2.

```



```

QY 201 TGENSEPLYKYPTDTGSGNAYLVNDYVYNNWKNGAPAEKLIIVGPTTYGHNFIILSNPNTGI 260
DB 61 TGENSEPLYKYPTDTGSGNAYLVNDYVYNNWKNGAPAEKLIIVGPTTYGHNFIILSNPNTGI 120
QY 261 GAPTSGAGPAGYAKESGIWAYEICTFLKNGATQGDAPQSVPPYAYQGNVWGVYDNIS 320
DB 121 GAPTSGAGPAGYAKESGIWAYEICTFLKNGATQGDAPQSVPPYAYQGNVWGVYDNIS 180
QY 321 FDIKAQWLKHNKFGGAMVWAIIDLDFTGTFGFCNOGKFPPLISTLKALGLQASACTAPAOPI 380
DB 181 FDIKAQWLKHNKFGGAMVWAIIDLDFTGTFGFCNOGKFPPLISTLKALGLQASACTAPAOPI 240
QY 381 EBITAAPSGSGSGSGSGSGSGSGGFCACVRANGLYPVANNRNFVHVCVNGVTVYQNCQ 440
DB 241 EBITAAPSGSGSGSGSGSGSGSGGFCACVRANGLYPVANNRNFVHVCVNGVTVYQNCQ 300
QY 441 AGLVFDTCDCCNWA 455
DB 301 AGLVFDTCDCCNWA 315

RESULT 8
Q31XA9 PRELIMINARY; PRT; 365 AA.
AC Q91XA9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to eosinophil chemotactic cytokine.
GN CHIA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE-SALIVARY GLAND;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011134; AAH11134.1; -.
DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM 14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN.1.
SQ SEQUENCE 365 AA; 39968 MW; C900BE0C4416F1DD CRC64;

Query Match 66.7%; Score 1670.5; DB 11; Length 365;
Best Local Similarity 80.4%; Pred. No. 7.9e-112;
Matches 295; Conservative 35; Mismatches 34; Indels 3; Gaps 1;

QY 88 MVSTPENRQTFTSVIKFLRQYEFGLDWDWYPSGRSPDPKHLFTVLVQEMREAEQ 147
DB 1 MVSTSONRQTFTSVIKFLRQYGFGLDWDWYPSGRSPDPKHLFTVLVQEMREAEQ 60
QY 148 BAKQINKPLMTAAVAGISINISQSYEIPQLSYLDYIHVNTYDLHGSWEYTGENSEPL 207
DB 61 EAGESNRPLMTAAVAGGISINISQSYEIPQLSYLDYIHVNTYDLHGSWEYTGENSEPL 120
QY 208 KYPTDTGSNAYLVNDYVYNNWKNGAPAEKLIIVGPTTYGHNFIILSNPNTGIGAPTS 267
DB 121 KYPTDTGSNAYLVNDYVYNNWKNGAPAEKLIIVGPTTYGHNFIILSNPNTGIGAPTS 180
QY 268 GPAGYAKESGIWAYEICTFLKNGATQGDAPQSVPPYAYQGNVWGVYDNISFIDKQW 327
DB 181 GPAGPYTQAGFWAYEICTFLRSGATEVWDASQSVPPYAYKANEWLYDNIKSFVSKAQW 240
QY 328 LKHNKFGGAMVWAIIDLDFTGTFGFCNOGKFPPLISTLKALGLQASACTAPAOPIEPI 387
DB 241 LKQNNFGGAMVWAIIDLDFTGTFGFCNOGKFPPLISTLKALGLQASACTAPAOPIEPI 300

```

```

QY 388 SSGSGSGSGSGSGSGSGGFCACVRANGLYPVANNRNFVHVCVNGVTVYQNCQAGLVFD 447
DB 301 ---GSGSGSGSGSGSGSGGFCADKADGLYPVADDRNAPQOCINGITTYQHCQAGLVFD 357
QY 448 SCDCCNW 454
DB 358 SCNCCNW 364

RESULT 9
Q35744 PRELIMINARY; PRT; 398 AA.
AC Q35744;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Secretary protein precursor.
GN CHI3L3 OR YMI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Chang N.C.A.; Hung S.I.; Hwa K.Y.; Kato I.; Chen J.E.; Liu C.H.;
RA Chang A.C.;
RT "A macrophage protein, Ymi, transiently expressed during inflammation
RT is a novel mammalian lectin.";
RL J. Biol. Chem. 276:17497-17506 (2001).
DR EMBL; M94584; AAB62394.2; -.
DR MGD; MGI:1330860; Chi3l3.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 398 SECRETORY PROTEIN.
SQ SEQUENCE 398 AA; 44458 MW; C1187661B99D1D1 CRC64;

Query Match 56.5%; Score 1415; DB 11; Length 398;
Best Local Similarity 68.1%; Pred. No. 1.8e-93;
Matches 256; Conservative 48; Mismatches 72; Indels 0; Gaps 0;

QY 1 YOLTCYFTNWAQYRPLGREFMPDNIDPCLCTHLYAFAGRONNEITTIEMNDVTLYQAFN 60
DB 22 YOLMCYFTSWAKDRPIEGSFKPGNIDPCLCTHLYAFAGRONNEITYTHEQDLRDYEALN 81
QY 61 GLKKNNSQLKTLTLLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFGLDWDWY 120
DB 82 GLKKNNSQLKTLTLLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFGLDWDWY 141
QY 121 PGRSGSPDQKHLFTVLVQEMREAEQBAKQINKPLMTAAVAGISINISQSYEIPQLS 180
DB 142 PGRSGSPDQKHLFTVLVQEMREAEQBAKQINKPLMTAAVAGISINISQSYEIPQLS 201
QY 181 QYLDYIHVNTYDLHGSWEYTGENSEPLKYPTDTGSNAYLVNDYVYNNWKNGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDKDGYTGENSEPLKYPTDTGSNAYLVNDYVYNNWKNGAPAEKLI 261
QY 241 VGFPYTHNFILSNPNTGIGAPTSAGPAGYAKESGIWAYEICTFLKNGATQGDAP 300
DB 262 VGFPYTHNFILSNPNTGIGAPTSAGPAGYAKESGIWAYEICTFLKNGATQGDAP 321
QY 301 QEVPIYQGNVWGVYDNISFIDKQWLNKHNKFGGAMVWAIIDLDFTGTFGFCNOGKFP 360
DB 322 QEVPIYQGNVWGVYDNISFIDKQWLNKHNKFGGAMVWAIIDLDFTGTFGFCNOGKFP 381
QY 361 TLKALGLQASACTAP 376
DB 382 TLKGLDNIHSASCRGP 397

```

```
RESULT 10
P70201
ID P70201 PRELIMINARY; PRT; 398 AA.
AC P70201;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-MAR-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2002 (TrEMBLrel. 20, Last annotation update)
DE ECF-L.
GN CHI3L3 OR MECF-L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=20092902; PubMed=10625674;
RA Ohashi M., Arita H., Hayai N.;
RT "Identification of a novel eosinophil chemotactic cytokine (ECF-L) as
RT a chitinase family protein.";
RL J. Biol. Chem. 275:1279-1286(2000).
DR EMBL; D87757; BAA13458.2; -
DR MGD; MG1:1330860; Chi3l3.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR P70201;
SQ SEQUENCE 398 AA; 4448 MW; 9D133FC490096793 CRC64;

Query Match 56.1%; Score 1407; DB 11; Length 398;
Best Local Similarity 67.8%; Pred. No. 6.8e-93;
Matches 255; Conservative 48; Mismatches 73; Indels 0; Gaps 0;

QY 1 YOLTCYFTNWAQYRGLGRFMPNDIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 60
DB 22 YQLMCYTTSWAKDRPTEGSKPGNIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 81
QY 61 GLKKNKSQKTLAIGGNWFGTAPFTAMVSTPNROTFTITSVKLFQYEFQGLDWDWEY 120
DB 82 GLKKNKTELKTLAIGGNWFGTAPFTAMVSTPNROTFTITSVKLFQYEFQGLDWDWEY 141
QY 121 PGSRGSPQDKHLFTVLVQEMREAFQEAQKQINKPRLMTAAVAAGISNIQSGYEPQLS 180
DB 142 PGSRGSPQDKHLFTVLVQEMREAFQEAQKQINKPRLMTAAVAAGISNIQSGYEPQLS 201
QY 181 QVLDIHWTYDLHGSWEGYTGENSEPLKYKPTDTSNAYLVNDYVMWYKDKGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKDGTYGENSEPLKYKPTDTSNAYLVNDYVMWYKDKGAPAEKLI 261
QY 241 VGFPTYGHNFTLSPNNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGDAP 300
DB 262 VGFPAHYHTFILDSPKNGIGDPTVSAGPPGKYTDESGLLAYEVCITFLNEGATEWDPAP 321
QY 301 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDDTGTFNCQKFPPLIS 360
DB 322 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDDTGTFNCQKFPPLIS 381
QY 361 TLKKAIGLOSASCTAP 376
DB 382 TLKGDNLNHSASCKGP 397

RESULT 11
Q91298
ID Q91298 PRELIMINARY; PRT; 398 AA.
AC Q91298;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Secreted protein precursor Ym2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=20092902; PubMed=10625674;
RA Ohashi M., Arita H., Hayai N.;
RT "Identification of a novel eosinophil chemotactic cytokine (ECF-L) as
RT a chitinase family protein.";
RL J. Biol. Chem. 275:1279-1286(2000).
DR EMBL; D87757; BAA13458.2; -
DR MGD; MG1:1330860; Chi3l3.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR P70201;
SQ SEQUENCE 398 AA; 4448 MW; 9D133FC490096793 CRC64;

Query Match 56.1%; Score 1407; DB 11; Length 398;
Best Local Similarity 67.8%; Pred. No. 6.8e-93;
Matches 255; Conservative 48; Mismatches 73; Indels 0; Gaps 0;

QY 1 YOLTCYFTNWAQYRGLGRFMPNDIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 60
DB 22 YQLMCYTTSWAKDRPTEGSKPGNIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 81
QY 61 GLKKNKSQKTLAIGGNWFGTAPFTAMVSTPNROTFTITSVKLFQYEFQGLDWDWEY 120
DB 82 GLKKNKTELKTLAIGGNWFGTAPFTAMVSTPNROTFTITSVKLFQYEFQGLDWDWEY 141
QY 121 PGSRGSPQDKHLFTVLVQEMREAFQEAQKQINKPRLMTAAVAAGISNIQSGYEPQLS 180
DB 142 PGSRGSPQDKHLFTVLVQEMREAFQEAQKQINKPRLMTAAVAAGISNIQSGYEPQLS 201
QY 181 QVLDIHWTYDLHGSWEGYTGENSEPLKYKPTDTSNAYLVNDYVMWYKDKGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKDGTYGENSEPLKYKPTDTSNAYLVNDYVMWYKDKGAPAEKLI 261
QY 241 VGFPTYGHNFTLSPNNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGDAP 300
DB 262 VGFPAHYHTFILDSPKNGIGDPTVSAGPPGKYTDESGLLAYEVCITFLNEGATEWDPAP 321
QY 301 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDDTGTFNCQKFPPLIS 360
DB 322 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDDTGTFNCQKFPPLIS 381
QY 361 TLKKAIGLOSASCTAP 376
DB 382 TLKGDNLNHSASCKGP 397
```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX PubMed=11553626;
RA Webb D.C., McKenzie A.N.J., Foster P.S.;
RT "Expression of the Ym2 lectin-binding Protein Is Dependent on
RT Interleukin (IL)-4 and IL-13 Signal Transduction. IDENTIFICATION OF A
RT NOVEL ALLERGY-ASSOCIATED PROTEIN.";
RL J. Biol. Chem. 276:41969-41976(2001).
DR EMBL; AF049765; AAL03953.1; -
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR P70201;
KW Signal.
FT SIGNAL.
SQ SEQUENCE 398 AA; 44515 MW; 79A74CASCDAA03850 CRC64;

Query Match 55.3%; Score 1387; DB 11; Length 398;
Best Local Similarity 66.5%; Pred. No. 1.8e-91;
Matches 250; Conservative 52; Mismatches 74; Indels 0; Gaps 0;

QY 1 YOLTCYFTNWAQYRGLGRFMPNDIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 60
DB 22 YQLMCYTTSWAKDRPTEGSKPGNIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 81
QY 61 GLKKNKSQKTLAIGGNWFGTAPFTAMVSTPNROTFTITSVKLFQYEFQGLDWDWEY 120
DB 82 GLKKNKTELKTLAIGGNWFGTAPFTAMVSTPNROTFTITSVKLFQYEFQGLDWDWEY 141
QY 121 PGSRGSPQDKHLFTVLVQEMREAFQEAQKQINKPRLMTAAVAAGISNIQSGYEPQLS 180
DB 142 PGSRGSPQDKHLFTVLVQEMREAFQEAQKQINKPRLMTAAVAAGISNIQSGYEPQLS 201
QY 181 QVLDIHWTYDLHGSWEGYTGENSEPLKYKPTDTSNAYLVNDYVMWYKDKGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKDGTYGENSEPLKYKPTDTSNAYLVNDYVMWYKDKGAPAEKLI 261
QY 241 VGFPTYGHNFTLSPNNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGDAP 300
DB 262 VGFPAHYHTFILDSPKNGIGDPTVSAGPPGKYTDESGLLAYEVCITFLNEGATEWDPAP 321
QY 301 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDDTGTFNCQKFPPLIS 360
DB 322 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDDTGTFNCQKFPPLIS 381
QY 361 TLKKAIGLOSASCTAP 376
DB 382 TLKRDNLNHSASCKGP 397

RESULT 12
Q8VHG1
ID Q8VHG1 PRELIMINARY; PRT; 402 AA.
AC Q8VHG1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ym2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RA Su W.B., Chang N.-C.A.;
RT "Genomic organization of the ym2 gene.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF461142; AAL66748.1; -
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
```

DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
SQ SEQUENCE 402 AA; 44975 MW; D935389319A59905 CRC64;

Query Match  
Best Local Similarity 55.2%; Score 1384; DB 11; Length 402;  
Matches 250; Conservative 52; Mismatches 73; Indels 0; Gaps 0;

QY 1 YQLTCYFTNWAQYRFGRLGFMFDPNDPCLCTHLIIYAFAGRONNEITTIEMNDVTLYQAFN 60  
DB 22 YQLMVCYTSWAKDRPTEGSFKPGNIDPCLCTHLIIYAFAGMKKNEITYLSEQDLRDYEALN 81

QY 61 GLKKNKSQLTKLLAIGGNWFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFWEY 120  
DB 82 GLKDRNTELTLLAIGCKWFGKPGAPFSSVSTPNRQTFTKSVIRELQYVNFDELNDWQY 141

QY 121 PGRSGSPQDKHLFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 180  
DB 142 PGRSGSPQDKHLFTVLVQEMRKAFEEESTLNLHPRLLLTSTGAGFIDVKSQYKIPELS 201

QY 181 QYLDYIHVMYTDLHGSWEGYTGENSEPLYKYPTDTGNSNAYLNVDYVMYWKONGAPAEKLI 240  
DB 202 QSLDVIQVMTYDLHDPKNGYTGENSEPLYKSPYDICKSADLNVDLSIITYWKOHGAASEKLI 261

QY 241 VGFPYAGNVMYGVYDNIKSFDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 360  
DB 301 QEVPIYAGNVMYGVYDNIKSFDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 360  
DB 322 QEVPIYAGNVMYGVYDNIKSFDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 381

QY 361 TLKKAALGLQSASCTA 375  
DB 382 TLKRDNLNVHSASCKA 396

RESULT 13  
QSVH43 PRELIMINARY; PRT; 402 AA.

AC QSVH43;  
DT 01-WAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Putative secretory protein.  
GN YM2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=STOMACH;  
RA Yang S.C., Su W.B.;  
RT "Tissue-specific Expression of Ym2 Protein."  
RL Submitted (DEC-2001) to the EMBL/GenBank/DDJ databases.  
DR EMBL; AY065557; AAL5751.1; -  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
SQ SEQUENCE 402 AA; 44985 MW; D935382E89A5981A CRC64;

Query Match  
Best Local Similarity 55.0%; Score 1379; DB 11; Length 402;  
Matches 249; Conservative 52; Mismatches 74; Indels 0; Gaps 0;

QY 1 YQLTCYFTNWAQYRFGRLGFMFDPNDPCLCTHLIIYAFAGRONNEITTIEMNDVTLYQAFN 60  
DB 22 YQLMVCYTSWAKDRPTEGSFKPGNIDPCLCTHLIIYAFAGMKKNEITYLSEQDLRDYEALN 81

QY 61 GLKKNKSQLTKLLAIGGNWFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFWEY 120  
DB 82 GLKDRNTELTLLAIGCKWFGKPGAPFSSVSTPNRQTFTKSVIRELQYVNFDELNDWQY 141

QY 121 PGRSGSPQDKHLFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 180  
DB 142 PGRSGSPQDKHLFTVLVQEMRKAFEEESTLNLHPRLLLTSTGAGFIDVKSQYKIPELS 201

QY 181 QYLDYIHVMYTDLHGSWEGYTGENSEPLYKYPTDTGNSNAYLNVDYVMYWKONGAPAEKLI 240  
DB 202 QSLDVIQVMTYDLHDPKNGYTGENSEPLYKSPYDICKSADLNVDLSIITYWKOHGAASEKLI 261

QY 241 VGFPYAGNVMYGVYDNIKSFDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 360  
DB 262 VGFPYAGNVMYGVYDNIKSFDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 360

QY 301 QEVPIYAGNVMYGVYDNIKSFDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 360  
DB 322 QEVPIYAGNVMYGVYDNIKSFDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 381

QY 361 TLKKAALGLQSASCTA 375  
DB 382 TLKRDNLNVHSASCKA 396

RESULT 14  
QSVH34 PRELIMINARY; PRT; 488 AA.

AC QSVH34;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Toad pancreatic chitinase (EC 3.2.1.14).  
GN TPCASE.  
OS Bufo japonicus (Japanese toad).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufonidae; Bufo.  
OX NCBI\_TaxID=8387;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PANCREAS;  
RA Oshima H.;  
RL Thesis (2001), Department of Department of Biology, Waseda University,  
RL Tokyo, Japan.  
DR EMBL; AJ345054; CAC87888.1; -  
DR InterPro; IPR002086; Aldenhyde\_dehydr.  
DR InterPro; IPR001579; Chitinase 18/2.  
DR InterPro; IPR002557; Chitin bind PerA.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF01607; CBM\_14; 1.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR PROSITE; PS00070; ALDEHYDE DEHYDR CYS; UNKNOWN\_1.  
DR PROSITE; PS01095; CHITINASE 18; UNKNOWN\_1.  
KW Glycosidase; Hydrolase; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 488 TOAD PANCREATIC CHITINASE.  
SQ SEQUENCE 488 AA; 54319 MW; E75E1AD3CB2B4919 CRC64;

Query Match  
Best Local Similarity 53.5%; Score 1340.5; DB 13; Length 488;  
Matches 249; Conservative 58; Mismatches 133; Indels 43; Gaps 4;

QY 1 YQLTCYFTNWAQYRFGRLGFMFDPNDPCLCTHLIIYAFAGRONNEITTIEMNDVTLYQAFN 60  
DB 19 YKLVCFYFTNWSQYRPGDQGYVPGNIDPCLCTHLVYAFATMNEHKIAPYEWNDVLYKQFN 78

QY 61 GLKKNKSQLTKLLAIGGNWFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFWEY 120  
DB 79 DLKQKNKLVTLAIGGNWFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFWEY 138

QY 121 PGRSGSPQDKHLFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 180  
DB 139 PGRSGSPQDKHLFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 198

QY 181 QYLDYIHVMYTDLHGSWEGYTGENSEPLYKYPTDTGNSNAYLNVDYVMYWKONGAPAEKLI 240  
DB 181 QYLDYIHVMYTDLHGSWEGYTGENSEPLYKYPTDTGNSNAYLNVDYVMYWKONGAPAEKLI 240

Db 199 QLLDFISVMTYDFHGWDTQSGHNSPLCKSGSTDYGLQYNIHFAMVYWKNGAPAEKLL 258  
QY 241 VGPPTTGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGWDA 300  
Db 259 LGPPTTGRTRFRNPNNDGVIPSGAGSAGPYTREAGFWAYVEICTWL-SGSTVKWIPD 317  
QY 301 QEVPYAYQGVVGYDNIKSFDIKAQWLKHNFKFGAMVAIDLDDFTGTFCNQGKPEPLIS 360  
Db 318 QRPVYACKSNEWGFNDQESYECKVFLKESGFGAMVAIDLDDFTGTFCNQGKPEPLIN 377  
QY 361 TLKAL-----GLQ-----SASCTAPAOPIETPITAAPSSGS 391  
Db 378 HLKSLLEGSTVNCBICGIGISITPTSATTTTTTAKPDCTTPEPPVTPPVPV--- 434  
QY 392 NGSSESSSGSGSGGFCVAVRANGLYPVANNRFAWHCVNGVTYQQNCOAGLVDFDSCDC 451  
Db 435 -----IDVDPNFCVEKTDGLHVNPLNTKFKYICANGRTYSMKCADGLVFOASCNC 484  
QY 452 CNW 454  
Db 485 CNW 487

## RESULT 15

Q13231  
ID Q13231 PRELIMINARY; PRT; 466 AA.  
AC Q13231;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Chitotriosidase precursor.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96064695; PubMed=7592832;  
RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;  
RT "Cloning of a cDNA encoding chitotriosidase, a human chitinase  
produced by macrophages.";  
RL J. Biol. Chem. 270:26252-26256(1995).  
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
CC EMBL; U29615; AAC50246.1; -;  
DR InterPro; IPR001579; Chitinase\_18/2.  
DR InterPro; IPR002557; Chitin\_bind\_Pera.  
DR InterPro; IPR001223; Glyco\_Hydro\_18.  
DR Pfam; PF01607; CBM\_14; 1.  
DR ProDom; PD00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD00471; Glyco\_hydro\_18; 1.  
DR SMART; SM00494; ChitB2; 1.  
DR PROSITE; PS01095; CHITINASE\_18; 1.  
DR Glycosidase; Hydrolase; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 466 CHITOTRIOSIDASE.  
SQ SEQUENCE 466 AA; 51681 MW; B4312D1E8895E386D CRC64;

Query Match 52.8%; Score 1323; DB 4; Length 466;

Best Local Similarity 52.2%; Pred. No. 8.7e-87;

Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 QLTCTFTNWAQYRPGIGRPMONIDPCLCTHLIYAFAGRNNEITTEWNDVTLQAFNG 61  
Db 23 KLVCYFTNWAQYRQGEARFLPKDPLSLCTHLIYAFAGTNNHOLSTEWNDVTLQAFNG 82  
QY 62 LKKNLSQKTLAIGGNFTGAPFTAMVSTPENRQFTTSVIKFLRQYFDFGLDFDWEYP 121  
Db 83 LKKNPFLKTLAIGGNFTGAPFTAMVSTPENRQFTTSVIKFLRQYFDFGLDFDWEYP 142  
QY 122 GSGSPQDKHLFTVLVQEMREAFQEAQKINPRLMVTAAVAGISNTQSGYEIPQLSQ 181  
Db 143 GSGSPAVDKERTFTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVYDAGYVDKIAQ 202

QY 182 YLDYIHVMTYDLHGSWEGVTGNSPLYKYPTDTGSNAVLYNDVYVNVYKNGCAPAEKLTIV 241  
Db 203 NUDFVNLWAYDFHGSWEKVTGNSPLYKROBESGAAASLNVDAAYQOHLQKTPASKLIL 262  
QY 242 GPPTYGHNFIILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGWDA 301  
Db 263 GWPTYGRSFTLASSSDTRVGAPATSGTGPPTKEGGLAYVEVCSW--KGATKQRIODQ 320  
QY 302 EVPYAYQGVVGYDNIKSFDIKAQWLKHNFKFGAMVAIDLDDFTGTFCNQGKPEPLIST 361  
Db 321 KVPYIFRDQNWGFODVESFKTKVSYLKQKGLGGAMVWALDLDLDFAGFSCNQGRYPLIQ 380  
QY 362 LKAL-----GLQASCTAPAOPIETPITAAPSGNGSGSSSGSGSGGFCVAVRAN 414  
Db 381 LRQELSLPYLPSGTPELEVPKPGQSEPP-----EKGPSFGQDTFCQKAD 425  
QY 415 GLYPVANNRNFHCHVCVNGVTYQQNCOAGLVDFDSCDCNW 454  
Db 426 GLYPNPRERSFYSACAAGRLFOQSCPTGLVFSNSCKCCTW 465

Search completed: June 29, 2003, 21:01:48  
Job time : 26.5769 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:57:59 ; Search time 11.7788 Seconds  
(without alignments)  
1136.565 Million cell updates/sec

Title: US-10-004-219b-14  
Perfect score: 2506  
Sequence: 1 YQLTCYFTNWAQYRGLGRF.....QQNCOAGLVFDTSCDCNWA 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/aaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/aaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/aaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/aaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/aaa/FCtUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/aaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1323	52.8	466	2	US-08-486-839-4
2	1323	52.8	466	3	US-09-151-011-4
3	1323	52.8	466	4	US-09-039-198A-2
4	1323	52.8	466	4	US-09-143-623-4
5	1323	52.8	466	4	US-08-877-599-2
6	1323	52.8	466	4	US-09-267-574-2
7	1317	52.6	466	4	US-09-039-198A-4
8	1317	52.6	466	4	US-08-877-599-4
9	1317	52.6	466	4	US-09-267-574-4
10	1177	47.0	373	4	US-09-039-198A-14
11	1177	47.0	373	4	US-09-039-198A-15
12	1177	47.0	373	4	US-08-877-599-14
13	1177	47.0	373	4	US-08-877-599-15
14	1177	47.0	373	4	US-09-267-574-14
15	1177	47.0	373	4	US-09-267-574-15
16	1173	46.8	387	3	US-08-486-839-6
17	1173	46.8	387	3	US-09-151-011-6
18	1172	46.8	387	4	US-09-343-623-6
19	970.5	38.7	385	2	US-08-694-915-2
20	970.5	38.7	416	2	US-08-694-915-4
21	968.5	38.6	383	4	US-09-459-749D-17
22	863	34.4	554	2	US-08-524-051-2
23	863	34.4	554	3	US-09-052-778-16
24	798.5	31.9	559	4	US-09-545-814-14
25	798.5	31.9	583	4	US-09-545-814-5
26	798.5	31.9	583	4	US-09-545-814-5
27	794.5	31.7	635	4	US-09-545-814-32

28	711.5	28.4	490	4	US-09-292-225-41	Sequence 41, Appl
29	711.5	28.4	509	4	US-09-292-225-35	Sequence 35, Appl
30	711.5	28.4	509	4	US-09-292-225-38	Sequence 38, Appl
31	708.5	28.3	536	4	US-09-292-225-21	Sequence 21, Appl
32	708.5	28.3	555	4	US-09-292-225-15	Sequence 15, Appl
33	708.5	28.3	555	4	US-09-292-225-18	Sequence 18, Appl
34	672	26.8	489	4	US-09-545-814-29	Sequence 29, Appl
35	472.5	18.9	389	1	US-07-939-501A-1	Sequence 1, Appl
36	472.5	18.9	389	4	US-08-448-398-7	Sequence 7, Appl
37	472.5	18.9	423	1	US-07-939-501A-10	Sequence 10, Appl
38	472.5	18.9	423	1	US-07-939-501A-12	Sequence 12, Appl
39	463.5	18.5	424	1	US-08-045-269C-2	Sequence 2, Appl
40	463.5	18.5	424	3	US-08-371-680-2	Sequence 2, Appl
41	462.5	18.5	424	5	PCT-US94-01198-2	Sequence 2, Appl
42	454	18.1	442	3	US-09-052-778-2	Sequence 2, Appl
43	402	16.0	371	2	US-08-591-629-8	Sequence 2, Appl
44	392.5	15.7	377	2	US-08-591-629-8	Sequence 8, Appl
45	388	15.5	452	3	US-09-052-778-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-08-486-839-4  
; Sequence 4, Application US/08486839  
; Patent No. 5928928  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A human chitinase, its recombinant  
; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11758  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; SOFTWARE:

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486.839  
FILING DATE: 07 - June - 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Baron, Ronald J.  
REGISTRATION NUMBER: 29,281  
REFERENCE/DOCKET NUMBER: 294-26  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-486-839-4

Query Match 52.8%; Score 1323; DB 2; Length 466;  
Best Local Similarity 52.2%; Pred. No. 1.4e-111;  
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;  
Query 2 QUTCYFTNWAQYRGLGRFMPDNIDPCLTHLIYAFAGRNNEITTEWVTLQAFNG 61  
Db 23 KLVCYFTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLYQEFNG 82

Query Match 52.8%; Score 1323; DB 3; Length 466;  
Best Local Similarity 52.2%; Pred. No. 1.4e-111;  
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 62 LKKNKSLKTLAIGWNGFTAPFTAMVSTPENRQTFTSVIKFLROYBFDGLDFDWEYP 121  
DB 83 LKKNKSLKTLAIGWNGFTAPFTAMVSTPENRQTFTSVIKFLROYBFDGLDFDWEYP 142

QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSIEIPOLSQ 181  
DB 143 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGRLLLSAAVPAQOTYVDAGYVDKIAQ 202

QY 182 YLDYIHVMTYDLHSGWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKDNKAPAEKLI 241  
DB 203 NLDYFVNLMAIDYFHSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKDNKAPAEKLI 262

QY 242 GFTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATOGWDAQ 301  
DB 263 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWS--KGATKORI 320

QY 302 EYPYAYQGNVWYGYDNIKSFDIKAKQWLNKHNKFGAMVWALDLDFTGTCNQKFP 361  
DB 321 KVPYIFRDQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDFTGTCNQKFP 380

QY 362 LKAL-----GLOSASCTAPAOPIEPIITAAPSGSGSGSSSGSGSGSGGFC 414  
DB 381 LRQELSLPLPSGTPELEVPKQPSEP-----EHGSPSPQDFTFCQKAD 425

QY 415 GLYPVANNRNFHVCNGVTVQONCOAGLVFDTSCDCNNW 454  
DB 426 GLYPNRRSSFSYCAAGRLFOQSCPTGLVFSNCKCCTW 465

## RESULT 2

US-09-151-011-4  
; Sequence 4, Application US/09151011  
; Patent No. 6057142  
; GENERAL INFORMATION:  
; APPLICANT: Hoffmann & Baron, LLP  
; CITY: Syosset  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11791  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; SOFTWARE:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/151,011  
; FILING DATE: 10 - September - 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Morris, Robert C.  
; REGISTRATION NUMBER: 42,910  
; REFERENCE/DOCKET NUMBER: 294-32 DIV  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3502  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; US-09-151-011-4

Query Match 52.8%; Score 1323; DB 3; Length 466;  
Best Local Similarity 52.2%; Pred. No. 1.4e-111;  
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 QLTCTFTWQVPRGLGRPMNDIPCLCTHLIYAFAGRONNEITTIWVTLVQAFNG 61  
DB 23 KLVCVFTWQVPRGLGRPMNDIPCLCTHLIYAFAGRONNEITTIWVTLVQAFNG 82

QY 62 LKKNKSLKTLAIGWNGFTAPFTAMVSTPENRQTFTSVIKFLROYBFDGLDFDWEYP 121  
DB 83 LKKNKSLKTLAIGWNGFTAPFTAMVSTPENRQTFTSVIKFLROYBFDGLDFDWEYP 142

QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSIEIPOLSQ 181  
DB 143 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGRLLLSAAVPAQOTYVDAGYVDKIAQ 202

QY 182 YLDYIHVMTYDLHSGWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKDNKAPAEKLI 241  
DB 203 NLDYFVNLMAIDYFHSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKDNKAPAEKLI 262

QY 242 GFTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATOGWDAQ 301  
DB 263 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWS--KGATKORI 320

QY 302 EYPYAYQGNVWYGYDNIKSFDIKAKQWLNKHNKFGAMVWALDLDFTGTCNQKFP 361  
DB 321 KVPYIFRDQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDFTGTCNQKFP 380

QY 362 LKAL-----GLOSASCTAPAOPIEPIITAAPSGSGSGSSSGSGSGSGGFC 414  
DB 381 LRQELSLPLPSGTPELEVPKQPSEP-----EHGSPSPQDFTFCQKAD 425

QY 415 GLYPVANNRNFHVCNGVTVQONCOAGLVFDTSCDCNNW 454  
DB 426 GLYPNRRSSFSYCAAGRLFOQSCPTGLVFSNCKCCTW 465

## RESULT 3

US-09-039-198A-2  
; Sequence 2, Application US/09039198A  
; Patent No. 6200951  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,198A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/34391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid



```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-039-198A-2

Query Match      52.8%; Score 1323; DB 4; Length 466;
Best Local Similarity 52.2%; Pred. No. 1.4e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 QUTCYFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGQNNNEITTEWNDVTLQAFNG 61
DB 23 KLVCFYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGTNNHQLSTTEWNETLQYFNG 82
QY 62 LKVNKSQLKTLLAIGGNWFGTAPFTAMVSTPENROTFTITSVKFLROYEFDGLDFDWEYP 121
DB 83 LKKMNPCLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 142
QY 122 GSRGSPPODKHLFTVLVOEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSGYEIPQLSQ 181
DB 143 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAQOTYVDAGYEVDKIAQ 202
QY 182 YLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTSNAYLNVDYVMNYKONGAPAEKLIIV 241
DB 203 NLDVFNLMAYDFHGSWEKVTGHSNPLYKQESGAASLNVDAAVQOVLQKTPASKLIL 262
QY 242 GPFTYGHNFILSNPTGIGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 263 GMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGMLAYEVCWS--KGATKQRIQDQ 320
QY 302 EYPYAYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTFCNOGKPELIST 361
DB 321 KVPYIFRDNQWVGDFDVSFETKVSYLKQGLGGAMVWALDLDDFAGFSCNOGRYPLIOT 380
QY 362 LKAL-----GLQASACTAPAEPIETAPSGSGNGSGSSSGSGSGSGGFCACAVRAN 414
DB 381 LKQELSLPLPSGTPELEVPKQGPSEP-----EHPSPGQDTFCQGRAD 425
QY 415 GLYPVANNRNPFWHCUNGVTYQONCOAGLVFDTSCDCCNW 454
DB 426 GLYPNPRSSSFYSCAAGRLFQOCSCTGLVFSNSCKCCTW 465

RESULT 4
US-09-343-623-4
; Sequence 4, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281

; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-343-623-4

Query Match      52.8%; Score 1323; DB 4; Length 466;
Best Local Similarity 52.2%; Pred. No. 1.4e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 QUTCYFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGQNNNEITTEWNDVTLQAFNG 61
DB 23 KLVCFYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGTNNHQLSTTEWNETLQYFNG 82
QY 62 LKVNKSQLKTLLAIGGNWFGTAPFTAMVSTPENROTFTITSVKFLROYEFDGLDFDWEYP 121
DB 83 LKKMNPCLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 142
QY 122 GSRGSPPODKHLFTVLVOEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSGYEIPQLSQ 181
DB 143 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAQOTYVDAGYEVDKIAQ 202
QY 182 YLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTSNAYLNVDYVMNYKONGAPAEKLIIV 241
DB 203 NLDVFNLMAYDFHGSWEKVTGHSNPLYKQESGAASLNVDAAVQOVLQKTPASKLIL 262
QY 242 GPFTYGHNFILSNPTGIGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 263 GMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGMLAYEVCWS--KGATKQRIQDQ 320
QY 302 EYPYAYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTFCNOGKPELIST 361
DB 321 KVPYIFRDNQWVGDFDVSFETKVSYLKQGLGGAMVWALDLDDFAGFSCNOGRYPLIOT 380
QY 362 LKAL-----GLQASACTAPAEPIETAPSGSGNGSGSSSGSGSGGFCACAVRAN 414
DB 381 LKQELSLPLPSGTPELEVPKQGPSEP-----EHPSPGQDTFCQGRAD 425
QY 415 GLYPVANNRNPFWHCUNGVTYQONCOAGLVFDTSCDCCNW 454
DB 426 GLYPNPRSSSFYSCAAGRLFQOCSCTGLVFSNSCKCCTW 465

RESULT 5
US-08-877-599-2
; Sequence 2, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
```

```

; SEQ ID NO 2
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-267-574-2

Query Match          52.8%; Score 1323; DB 4; Length 466;
Best Local Similarity 52.2%; Pred. No. 1.4e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY      2   QLTCTYTNAQAQPGRLGRFMPONIDPCLCTHLIYAIPAGRONNEITTIWENDVLYQAIFNG 61
DB      2   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 82
QY      62   LKKNLSQLTKLLAIGWNFTGTPATTAMVTSPENRQTITSVKFLROYEPDGLDFWEYP 121
DB      62   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 142
QY      83   LKKNNPKLTKLLAIGWNFTGTOKFTDMVATANNRQTFVNSAIRFLKYSPDGLDLWEYP 142
DB      83   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181
QY     122   GSRGPSPODKHFTVLVQBMRAFEAEQAIKNKPLMWTAATAVAAGISNIOQSVEIPOLSQ 181
DB     122   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 202
QY     143   GSGGSFAVDKERFTTLVODLANAFQEATSGKERLLLSANVPAGTYVDAGYEVDKIAQ 202
DB     143   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 241
QY     182   YLDYIHVMYTDLHGSEWGVTGENSPPLYKYPTDTGSNAYILNVDMYNWKNDGAPASKLIV 241
DB     182   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 262
QY     203   NLDFFNLMAFYDFHGSWEKVTHGNSPLYKREQESSGAASLNVDAAVQQWLQKGTPASKLIL 262
DB     203   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 301
QY     242   GPPTYGHNTILNSPNSTGTTGAPTSGAGPAGPYAKESGIWAYIYEICTFLKXGATQGWDAPQ 301
DB     242   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 320
QY     263   GMPTYGRSFTLASSSDTRVGAAFGATGPGTGPFTKEGGMLLAYIVECSW--KGATKQRIQDQ 320
DB     263   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 361
QY     302   EVPNAYQGNVWGYDNIKSFDIKAOWLKHNFKFGAMVAIDLDDFTGTCNQCKFPPLIST 361
DB     302   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 380
QY     321   KVPYIIFRDNQWGFDDVESFKTKVSYLKQKGLGGAMVALDDLDPAGFSCNQGRYFLIQ 380
DB     321   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 414
QY     362   LKAL-----GLQASACTAPAQPIETITAAPSNGCNGSSSSSGSGSGSGFCVAVRAN 414
DB     362   ::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||: 425
QY     381   LRQELSLPLPSGTPELVKPGQPSEP-----EHGPPSPQDFTFCQKAD 425
DB     381   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 454
QY     415   GLYPVANNNNAFWHCVNGVTVYQNCOAGLVFDTSDCCCNW 454
DB     415   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 465
QY     426   GLYPNPRERSSPYSCAAGRLFQOCSPTGLVFNSCKCCTW 465
DB     426   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
US-09-039-198A-4
; Sequence 4, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-039-198A-4

Query Match 52.6%; Score 1317; DB 4; Length 466;  
Best Local Similarity 52.0%; Pred. No. 4.8e-111;  
Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;

```
QY 2 QLTCTFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIENNDVTLQAFNG 61
DB 23 KLVCFYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGWTHQLSTTEWNDLTYQEFNG 82
QY 62 LKNKNSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVIKFLROYEFDGLDFDWEYP 121
DB 83 LKKNPPLKTLAIGWNFSTQKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 142
QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNIQSGYEIPQLSQ 181
DB 143 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGRLLLSAAVPAQGTYYDAGYEVDKIAQ 202
QY 182 YLDYIHVMYDHLHSGWEGVTGENSPLYKYPTDTGNSVNLVNDVYVWYKNDGAPAEKLIV 241
DB 203 NLDFVNLMAIDPHGSEKVTGHSPLRYKQESGAASLNVDAAVQOQLKGTPTASKLIL 262
QY 242 GPPTYGHNFIPLSNPTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGWDAPO 301
DB 263 GMPTYGSRFTLASSSDTRVGATSGTGPPTKEGMLAYEVCWS--KGATKORIQDO 320
QY 302 EYPYAYQGNVWVGYDNIKSFDIKAQWLKHNFKGGAMVWALDLDFTGTFCNGKPEPLIST 361
DB 321 KVPYIFRDNQWVGDFDVESFKTKVSYLKQKGLGGAMVWALDLDPAFSGSCNOGRYPLIQT 380
QY 362 LKAL-----GLQASCTAPAEPIETAPSGSGSGSSSSSGSGSGSGGSCGFCAVRAN 414
DB 381 LKQELSLPPLSGTPELEVPKQPSEP-----EHGPPSGQDTFCQGRAD 425
QY 415 GLYPVANNRNFHWCNVGTYQONCOAGLVFDTSCDCCNW 454
DB 426 GLYPNPRSSSFYSCAAGRLFQOCSCTGLVFSNCKCCTW 465
```

## RESULT 8

US-08-877-599-4  
Sequence 4, Application US/08877599  
Patent No. 6372212

## GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.  
TITLE OF INVENTION: Chitinase Materials and Methods  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/877,599  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/663,618

FILING DATE: 14-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 27866/33994  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-877-599-4

Query Match 52.6%; Score 1317; DB 4; Length 466;  
Best Local Similarity 52.0%; Pred. No. 4.8e-111;  
Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;

```
QY 2 QLTCTFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIENNDVTLQAFNG 61
DB 23 KLVCFYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGWTHQLSTTEWNDLTYQEFNG 82
QY 62 LKNKNSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVIKFLROYEFDGLDFDWEYP 121
DB 83 LKKNPPLKTLAIGWNFSTQKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDLDWEYP 142
QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNIQSGYEIPQLSQ 181
DB 143 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGRLLLSAAVPAQGTYYDAGYEVDKIAQ 202
QY 182 YLDYIHVMYDHLHSGWEGVTGENSPLYKYPTDTGNSVNLVNDVYVWYKNDGAPAEKLIV 241
DB 203 NLDFVNLMAIDPHGSEKVTGHSPLRYKQESGAASLNVDAAVQOQLKGTPTASKLIL 262
QY 242 GPPTYGHNFIPLSNPTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGWDAPO 301
DB 263 GMPTYGSRFTLASSSDTRVGATSGTGPPTKEGMLAYEVCWS--KGATKORIQDO 320
QY 302 EYPYAYQGNVWVGYDNIKSFDIKAQWLKHNFKGGAMVWALDLDFTGTFCNGKPEPLIST 361
DB 321 KVPYIFRDNQWVGDFDVESFKTKVSYLKQKGLGGAMVWALDLDPAFSGSCNOGRYPLIQT 380
QY 362 LKAL-----GLQASCTAPAEPIETAPSGSGSGSSSSSGSGSGSGGSCGFCAVRAN 414
DB 381 LKQELSLPPLSGTPELEVPKQPSEP-----EHGPPSGQDTFCQGRAD 425
QY 415 GLYPVANNRNFHWCNVGTYQONCOAGLVFDTSCDCCNW 454
DB 426 GLYPNPRSSSFYSCAAGRLFQOCSCTGLVFSNCKCCTW 465
```

## RESULT 9

US-09-267-574-4  
Sequence 4, Application US/09267574  
Patent No. 6395571

## GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.  
APPLICANT: Tjoelker, Larry W.  
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
FILE REFERENCE: 27866/35407  
CURRENT APPLICATION NUMBER: US/09/267,574  
CURRENT FILING DATE: 1999-03-12  
EARLIER APPLICATION NUMBER: 09/039,198  
EARLIER FILING DATE: 1998-03-12  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 4  
LENGTH: 466  
TYPE: PRT  
ORGANISM: Homo sapiens

US-09-267-574-4

Query Match 52.6%; Score 1317; DB 4; Length 466;  
Best Local Similarity 52.0%; Pred. No. 4.8e-111;  
Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;

QY 2 QLTCTFTNNAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITIEWNDVLYQAFNG 61  
DB 23 KLVCYFTNNAQYRQGEARFLPKDLPSLCTHLIYAFAGTNHQLSTTEWDETLVQEFNG 82  
QY 62 LKKNKSQLKTLAIGWNFGTAPFTAMVSTPENROTFTSVIKFLROYEFDGLDFDWEYP 121  
DB 83 LKKNPKLTLAIGWNFGTAPFTAMVSTPENROTFTSVIKFLROYEFDGLDFDWEYP 142  
QY 122 GSRGSPQDKHLFTVLQEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSGYEIPQLSQ 181  
DB 143 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGERULLLSAAVPAQOTYVDAGYEVDKIAQ 202  
QY 182 YLDYIHVMTYDLHSGWEGTGENSPLYKYPTDTGNSNAYLNVDYVMYWKDNGAPAEKLI 241  
DB 203 NLDYFVNLMAVDYFHSGWEKVTGHNSPLYKQESGAAASLNVDAAVQOVLQKGTPTASKLIL 262  
QY 242 GPTYGHNFILSPNTGTGAPTSAGAPGAPYAKESGIWAYEICTFLKNGATQGWDAPO 301  
DB 263 GMPTGRSFLLASSSDTRVGAPATGSGTPTKGGMLAYEVCWS--KGATKQRIQDQ 320  
QY 302 EYPYAYQGNVWGYDNIKSFIDIAQWLKHNKFGAMWALDLDFTGTFCNQKRPPLIST 361  
DB 321 KVPYIFRDQWGVDFDVESEFKTKVSYLKQKGLGGAMWALDLDFTGTFCNQKRPPLIST 380  
QY 362 LKALGL 368  
DB 360 LRQELSL 366

RESULT 10

US-09-039-198A-14  
; Sequence 14, Application US/09039198A  
; Patent No. 6200951  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,198A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/34391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-039-198A-14

Query Match 47.0%; Score 1177; DB 4; Length 373;  
Best Local Similarity 56.9%; Pred. No. 1.7e-98;  
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

QY 2 QLTCTFTNNAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITIEWNDVLYQAFNG 61  
DB 2 KLVCYFTNNAQYRQGEARFLPKDLPSLCTHLIYAFAGTNHQLSTTEWDETLVQEFNG 61  
QY 62 LKKNKSQLKTLAIGWNFGTAPFTAMVSTPENROTFTSVIKFLROYEFDGLDFDWEYP 121  
DB 62 LKKNPKLTLAIGWNFGTAPFTAMVSTPENROTFTSVIKFLROYEFDGLDFDWEYP 121  
QY 122 GSRGSPQDKHLFTVLQEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSGYEIPQLSQ 181  
DB 122 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGERULLLSAAVPAQOTYVDAGYEVDKIAQ 181  
QY 182 YLDYIHVMTYDLHSGWEGTGENSPLYKYPTDTGNSNAYLNVDYVMYWKDNGAPAEKLI 241  
DB 182 NLDYFVNLMAVDYFHSGWEKVTGHNSPLYKQESGAAASLNVDAAVQOVLQKGTPTASKLIL 241  
QY 242 GPTYGHNFILSPNTGTGAPTSAGAPGAPYAKESGIWAYEICTFLKNGATQGWDAPO 301  
DB 242 GMPTGRSFLLASSSDTRVGAPATGSGTPTKGGMLAYEVCWS--KGATKQRIQDQ 299  
QY 302 EYPYAYQGNVWGYDNIKSFIDIAQWLKHNKFGAMWALDLDFTGTFCNQKRPPLIST 361  
DB 300 KVPYIFRDQWGVDFDVESEFKTKVSYLKQKGLGGAMWALDLDFTGTFCNQKRPPLIST 359  
QY 362 LKALGL 368  
DB 360 LRQELSL 366

RESULT 11

US-09-039-198A-15  
; Sequence 15, Application US/09039198A  
; Patent No. 6200951  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,198A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/34391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-039-198A-15

Query Match 47.0%; Score 1177; DB 4; Length 373;  
Best Local Similarity 56.9%; Pred. No. 1.7e-98;  
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

```
QY 2 QLTCTFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIENWDTLYQAFNG 61
DB 2 KLVCFYTNWAQYRQGEAREFLPKDLPSCSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFNG 61
QY 62 LKNKNSQLKTLTLLAIGWNFGTAPFTAMVSTPENROTFTTSVTKFLROYEFDGLDDEWYEP 121
DB 62 LKXMPKLTLLAIGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDDEWYEP 121
QY 122 GSRGSPPODKHLFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNQSVEIPLQSQ 181
DB 122 GSGSPAVDKERFTTLVQDLANAFQEAQTSCKERLLLSAAVPAQOTYVDAGYVDKIAQ 181
QY 182 YLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLNVDYVMYWKDNGAPAEKLIV 241
DB 182 NLDYFNLMAFDHGSWEKVTGHNSPLYKRBESGAAASLNVDAAVQOVLQKGTSPASKLIL 241
QY 242 GPFTYGHNFILSNPNTGIGAPTSGAGPAGPVAKESGIWAYYEICTFLKNGATQGWDAPO 301
DB 242 GMPTYGRSFTLASSDTRVGAPATSGTGPFTKGGMLAYEVCWSW--KGATKQRIQDQ 299
QY 302 EYPYAYQGVNVGVYDNIKSFDIKAKWLKHNKFGGAMVWALDLDFTGTFCNQKKEPLIST 361
DB 300 KVPYIFRDNQWGVDFDVESEFKTKVSYLKQKGLGGAMVWALDLDDEAGFSCNOGRYPILOT 359
QY 362 LKKAALGL 368
DB 360 LRQELSL 366
```

## RESULT 12

US-08-877-599-14  
Sequence 14, Application US/08877599  
Patent No. 6372212

## GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.  
TITLE OF INVENTION: Chitinase Materials and Methods  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/877,599  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/663,618  
FILING DATE: 14-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCES/DOCKET NUMBER: 27866/33994  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-877-599-14

```
Query Match 47.0%; Score 1177; DB 4; Length 373;  
Best Local Similarity 56.9%; Pred. No. 1.7e-98;  
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;
```

```
QY 2 QLTCTFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIENWDTLYQAFNG 61
DB 2 KLVCFYTNWAQYRQGEAREFLPKDLPSCSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFNG 61
QY 62 LKNKNSQLKTLTLLAIGWNFGTAPFTAMVSTPENROTFTTSVTKFLROYEFDGLDDEWYEP 121
DB 62 LKXMPKLTLLAIGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSFDFGLDDEWYEP 121
QY 122 GSRGSPPODKHLFTVLVQEMREAFQEAQKINPKRLMVTAAVAAGISNQSVEIPLQSQ 181
DB 122 GSGSPAVDKERFTTLVQDLANAFQEAQTSCKERLLLSAAVPAQOTYVDAGYVDKIAQ 181
QY 182 YLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLNVDYVMYWKDNGAPAEKLIV 241
DB 182 NLDYFNLMAFDHGSWEKVTGHNSPLYKRBESGAAASLNVDAAVQOVLQKGTSPASKLIL 241
QY 242 GPFTYGHNFILSNPNTGIGAPTSGAGPAGPVAKESGIWAYYEICTFLKNGATQGWDAPO 301
DB 242 GMPTYGRSFTLASSDTRVGAPATSGTGPFTKGGMLAYEVCWSW--KGATKQRIQDQ 299
QY 302 EYPYAYQGVNVGVYDNIKSFDIKAKWLKHNKFGGAMVWALDLDFTGTFCNQKKEPLIST 361
DB 300 KVPYIFRDNQWGVDFDVESEFKTKVSYLKQKGLGGAMVWALDLDDEAGFSCNOGRYPILOT 359
QY 362 LKKAALGL 368
DB 360 LRQELSL 366
```

## RESULT 13

US-08-877-599-15  
Sequence 15, Application US/08877599  
Patent No. 6372212

## GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.  
TITLE OF INVENTION: Chitinase Materials and Methods  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/877,599  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/663,618  
FILING DATE: 14-JUN-1996  
ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Heien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 27866/33994  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-877-599-15

Query Match 47.0%; Score 1177; DB 4; Length 373;  
Best Local Similarity 56.9%; Pred. No. 1.7e-98;  
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

```
QY 2 QLTCTFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIWNDVTLYQAFNG 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 KLVCTFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFNG 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 LKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTTSVIKFLRQYEFDFGLDFDWEYP 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 LKKNMPLKTLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEYP 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 GSRGSPPODKHLFTVLVOEMREAFQEAQKINPKRLMVTAAVAGISNIQSGEYEPQLSQ 181
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 GSQGSFAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAQQTVDAGYEVDKIAQ 181
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 YLDYIHVMYDYLHGSWEGYTGENSEPLYKYPTDGTGSNAYLNVDYVMNWKDNGAPAEKLI 241
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOMLQKGTGPASKLIL 241
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 GFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDAPO 301
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGGMLAYEVCWS--KGATKQRIQDQ 299
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 302 EYPYAYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTCNOGKFPPLIST 361
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 KVPYIFRDNQWGFDDVESFTKVSILKQKGLGGAMVWALDLDLDDPAGFSCNQGRYPLIQT 359
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 362 LKKALGL 368
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 LRQELSL 366
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

#### RESULT 14

US-09-267-574-14

Sequence 14, Application US/09267574

Patent No. 6395571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

APPLICANT: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

FILE REFERENCE: 27866/35407

CURRENT APPLICATION NUMBER: US/09/267,574

CURRENT FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: 09/039,198

EARLIER FILING DATE: 1998-03-12

NUMBER OF SEQ ID NOS: 39

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 14

LENGTH: 373

TYPE: PRT

ORGANISM: Homo sapiens

US-09-267-574-14

Query Match

Best Local Similarity 47.0%; Score 1177; DB 4; Length 373;

Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

```
QY 2 QLTCTFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIWNDVTLYQAFNG 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 KLVCTFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFNG 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 LKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTTSVIKFLRQYEFDFGLDFDWEYP 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 LKKNMPLKTLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEYP 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 GSRGSPPODKHLFTVLVOEMREAFQEAQKINPKRLMVTAAVAGISNIQSGEYEPQLSQ 181
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 GSQGSFAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAQQTVDAGYEVDKIAQ 181
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 YLDYIHVMYDYLHGSWEGYTGENSEPLYKYPTDGTGSNAYLNVDYVMNWKDNGAPAEKLI 241
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOMLQKGTGPASKLIL 241
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 GFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDAPO 301
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGGMLAYEVCWS--KGATKQRIQDQ 299
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 302 EYPYAYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTCNOGKFPPLIST 361
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 KVPYIFRDNQWGFDDVESFTKVSILKQKGLGGAMVWALDLDLDDPAGFSCNQGRYPLIQT 359
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 362 LKKALGL 368
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 LRQELSL 366
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

#### RESULT 15

US-09-267-574-15

Sequence 15, Application US/09267574

Patent No. 6395571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

APPLICANT: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

FILE REFERENCE: 27866/35407

CURRENT APPLICATION NUMBER: US/09/267,574

CURRENT FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: 09/039,198

EARLIER FILING DATE: 1998-03-12

NUMBER OF SEQ ID NOS: 39

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 15

LENGTH: 373

TYPE: PRT

ORGANISM: Homo sapiens

US-09-267-574-15

Query Match 47.0%; Score 1177; DB 4; Length 373;

Best Local Similarity 56.9%; Pred. No. 1.7e-98;

Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

```
QY 2 QLTCTFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIWNDVTLYQAFNG 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 KLVCTFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFNG 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 LKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTTSVIKFLRQYEFDFGLDFDWEYP 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 LKKNMPLKTLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEYP 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 GSRGSPPODKHLFTVLVOEMREAFQEAQKINPKRLMVTAAVAGISNIQSGEYEPQLSQ 181
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 GSQGSFAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAQQTVDAGYEVDKIAQ 181
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 YLDYIHVMYDYLHGSWEGYTGENSEPLYKYPTDGTGSNAYLNVDYVMNWKDNGAPAEKLI 241
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOMLQKGTGPASKLIL 241
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 GFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDAPO 301
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 242 GMPYGRSFTLASSDTRVGAPATCGTGPFTKEGMLAYEVCW--KGATKQRIODQ 299  
Qy 302 EYPYAYQGNVWYGYDNIKSFDIKAOWLKHNKFGGAMVWALDLDFTGTFCNOGKEPLIST 361  
Db 300 KVPYIFRDNQWVGFDVSEFVKVSYLKQKGLGGAMVWALDLDLDDPAGPSCNOGCRYPLIQT 359  
Qy 362 LKKALGL 368  
Db 360 LRQELSL 366

Search completed: June 29, 2003, 21:03:29  
Job time : 11.7788 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:03:33 ; Search time 2957.5 Seconds  
(without alignments)  
4477.355 Million cell updates/sec

Title: US-10-004-219B-14  
Perfect score: 2506  
Sequence: 1 YQLTCYFTNWAQYRPLGRF.....QQNCOAGLVFDTSCDCNWA 455

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DEV=xl  
-Q/cn2\_1/USPTO\_spool/US10004219/runat\_24062003\_160228\_8159/app\_query.fasta\_1.1877  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CN 1 1 6828 @runat\_24062003\_160228\_8159 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pin.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sv.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2506	100.0	1625	9	AF290004 Homo sapi
2	2137	85.3	1369	6	AX405989 Sequence
3	2119	84.6	1354	9	AB025008 Homo sapi
4	2097.5	83.7	1526	10	AF154571 Mus muscu
5	2097.5	83.7	1538	10	BC011134 Mus muscu
6	2097.5	83.7	1557	10	BC034548 Mus muscu
7	2089.5	83.4	1530	10	AF290003 Mus muscu
8	2039	81.4	1529	4	AB051629 Bos tauru
9	1768.5	70.6	1188	9	AB025009 Homo sapi
10	1742	69.5	4250	9	AK098814 Homo sapi
11	1415	56.5	1525	10	M94584 Mus musculu
12	1407	56.1	1506	10	D87757 Mus musculu
13	1384	55.2	1209	10	AY043765 Mus muscu
14	1379	55.0	1209	10	AY065557 Mus muscu
15	1362.5	54.4	91569	9	AL356387 Human DNA
16	1353	54.0	93740	9	AL513202 Human DNA
17	1340.5	53.5	1562	5	BJA345054 Bufo japo
18	1323	52.8	1633	9	HSU29615 Human chito
19	1323	52.8	1636	6	AR206041 Sequence
20	1323	52.8	1636	6	AR212138 Sequence
21	1323	52.8	1636	6	AR108750 Sequence
22	1323	52.8	1643	6	AR172107 Sequence
23	1317	52.6	1656	6	AR206042 Sequence
24	1317	52.6	1656	6	AR212139 Sequence
25	1317	52.6	1656	6	AX108752 Sequence
26	1311.5	52.3	1710	9	HSU62662 Homo sapien
27	1311.5	52.3	1713	6	AR172108 Sequence
28	1225.5	48.9	1599	9	AK055165 Homo sapi
29	1011.5	40.4	2504	6	E09046 cDNA encodi
30	1011.5	40.4	2525	10	MUSO3P Mouse mRNa
31	1010.5	40.3	6439	3	AB074977 Haemaphys
32	985.5	39.3	1741	9	HUMH33G Human glyco
33	983.5	39.2	1801	9	BC008568 Homo sapi
34	983	39.2	1281	4	AY081150 Capra hir
35	980	39.1	3946	3	AY051988 Drosophil
36	972	38.8	1733	4	SSU19300 Sus scrofa
37	972	38.8	1733	4	SSGFP38D S.scrofa 38
38	971.5	38.8	1149	6	E01500 cDNA encodi
39	971.5	38.8	1149	6	E01501 cDNA encodi
40	970.5	38.7	1418	9	HSU49835 Human YKL-3
41	970.5	38.7	1433	6	AR042834 Sequence
42	970.5	38.7	1434	9	HSU58314 Human chiti
43	970.5	38.7	1500	9	HSU58515 Human chiti
44	970.5	38.7	1526	6	AR042835 Sequence
45	966.5	38.6	1449	9	BC011460 Homo sapi

ALIGNMENTS

AF290004 1625 bp mRNA linear PRI 27-FEB-2001  
LOCUS Homo sapiens acidic mammalian chitinase precursor, mRNA, complete  
DEFINITION cds.  
ACCESSION AF290004  
VERSION AF290004.1 GI:12597292  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1625)  
AUTHORS Boot R.G., Blommaert, E.F., Swart, E., Chauharali-van der Vlugt, K.,  
Bi, J.N., Moe, C., Place, A. and Aerts, J.M.  
TITLE Identification of a novel acidic mammalian chitinase distinct from  
chitotriosidase  
J. Biol. Chem. 276 (9), 6770-6778 (2001)  
JOURNAL  
MEDLINE 21125893  
PUBMED 11085997  
REFERENCE 2 (bases 1 to 1625)  
AUTHORS Boot, R.G., Verhoeck, M., Swart, E. and Aerts, J.M.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical  
Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105  
AZ, The Netherlands  
FEATURES  
source  
1..1625  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="stomach; lung"  
104..1534  
/EC\_number="3.2.1.14"  
/notes="chitinase precursor; AMCase; chitinase activity  
demonstrated by expression in COS cells; similar to  
sequences presented in GenBank Accession Numbers AB025008  
and AB025009"  
/codon\_start=1  
/product="acidic mammalian chitinase precursor"  
/protein\_id="AAC60019.1"  
/db\_xref="GI:12597293"  
/translation="MTKLLILLGLVLILQLGSVALQVLCYFTNNWAQYRPGIGRMPD  
NDPCLCTHLIYAFAGQRNEITIEWNVTLTYQAFNGLNKNSQLKLLAIGWNFG  
TAPFTAMVSTPNROFTITSVKFLQYDFDGLDFWEYFGSRGSPQDKHLFTVLVQ  
ENREAEQAKINIKRLMTAAVAGISNIOSGVEIPOLSOVDYIHVMTVDLHGSW  
EGVTGNSPLYKPTDTGNAYLVNDVYVNVKNGAPAEKLIIVGPTTYGHNFILSNP  
SVTIGASPTSGAGPAGYAKESGIWAYIEICTFLENGATQGDWAQEVYPAVQGVV  
GDNISFDIKQWLKHNKFGGMVAIDDDFTGFCNQGRPLLIISTLUKALGLQSA  
SCTAPAPQIEPTAAPSGSGSGSSGSSGSGSGGFCVAVRANGLYPVANNRNFPHC  
VNGVTYQQNCQAGLVFDTSDCCNWA"  
sig\_peptide 104..166  
BASE COUNT 387 a 468 c 399 g 371 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,03e-184 Length: 1625  
Score: 2506.00 Matches: 455  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-004-219b-14 (1-455) x AF290004 (1-1625)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
Db 167 TACCAGCTGACATGCTACTTACCACTGGGCCAGTACCGGCAGCCCTGGGGCGCTTC 226  
QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40  
Db 227 ATGCCTGACACATCGACCCCTGCCTCTGTACCACCTGATCTACGCCCTTGTGGGAGG 286  
QY 41 GlnAsnAsnGluIleThrThrIleGluTyrAsnAspValThrLeuTyrGlnAlaPheAsn 60

287 CAGAAACACGAGATCACCACCATCGAATGGAACGATGTGACTCTCTACCAAGCTTTCAAT 346  
QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80  
Db 347 GGCCTGAAATAAAGAACAGCCAGCTGAAATCTCTCGGCCATTTGGAGGCTGGAACTTC 406  
QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
Db 407 GGGACTGCCCTTTTCACTGCCATGGTTTCTACTCTGAGAACCCGACAGCTTTTCATACC 466  
QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
Db 467 TCAGTCATCAATCTCTCGCCAGTATGAGTTTGAAGGCTGGAGCTTTGACTGGGAGTAC 526  
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
Db 527 CCTGGCTCTCGGGGAGCCCTCTCAGCAACAGCATCTCTTCTCCTGCTGGTGAGGAA 586  
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
Db 587 ATGCGTGAAGCTTTTGGAGGAGGAGCCAGCAGATCAACAAGCCAGGCTGATGGTCACT 646  
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
Db 647 GCTGCAGTAGCTGCTGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCACTGTCA 706  
QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
Db 707 CAGTACCTGGAGCTATCATCTGTCATGACCTCCATGGCTCTCTGGGAGGCTGAC 766  
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
Db 767 ACTGGAGAGAACAGCCCTCTTACAATAATCCGAGCTGACACCGCGCAGCAACCTTACCTC 826  
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240  
Db 827 AATGTGGATTATGTGATGAACCTACTCTGGAAGGACAAATGGAGCACCAGCTGAGAAGCTCATC 886  
QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260  
Db 887 GTTGGATTCCTTACCTATGAGCACAACTTCATCTCGAGCAACCCCTCCACACTGGAATT 946  
QY 261 GlyAlaProThrSerGlyValaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280  
Db 947 GGTGCCCCCACCCTCTGGTGTCTCTGCTGGGCCCTATGCCAAGAGTCTGGGATCTGG 1006  
QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300  
Db 1007 GCTTACTACGAGATCTGTACCTTCTGAAAAATGGAGCCACTCAGGGATGGGATGCCCT 1066  
QY 301 GlnGluValProTyrAlaTyrGlnGlyValAsnValTrpValGlyTyrAspAsnIleLysSer 320  
Db 1067 CAGGAAGTGCCCTTATGCTTATCAGGGCAATGTGGGTGGCTATGCAACATCAAGAGC 1126  
QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340  
Db 1127 TTCGATATTAAAGCTCAATGGCTTAAGCACACAATAATTTGGAGGCGCCATGCTGGGCC 1186  
QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360  
Db 1187 ATTGATCTGGATGACTTCTACGGCCTTCTGCAACACAGGCAAGTTTCCCCATACTCC 1246  
QY 361 ThrLeuLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380  
Db 1247 ACCCTGAAGAGGCCCTCGGCTTCAGAGTGCAGAGTGCAGGCTCCAGCTCAGGCCATT 1306  
QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400  
Db 1307 GAGCCAATAACTGCTGCTCCAGTGGCAGCGGGAACGGAGCGGAGTAGCAGCTCTGGA 1366  
QY 401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420  
Db 1367 GGCAGCTCGGAGGAGTGGATTCTGTGCTGTGAGAGCAACGGGCTCTTACCCCGTGGCA 1426

```
QY 421 AenAsnArgAsnAlaPheThrHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440
Db 1427 AATAACAGAAATGCTTGGACACTGCGTGAATGAGTCACTACAGCAGAACTGCCAG 1486
QY 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrpAla 455
Db 1487 GCGGGCTGTCTTCGACACCACTGTGATTGCTGCAACTGGCA 1531

RESULT 2
AX405989 1369 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 404 from Patent WO0222660.
DEFINITION AX405989
ACCESSION AX405989
VERSION AX405989.1 GI:21439417
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
Novel nucleic acids and polypeptides
Patent: WO 0222660-A 404 21-MAR-2002;
HYSBO, INC. (US)
FEATURES
source Location/Qualifiers
1..1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
154..1260
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35096.1"
/db_xref="GI:21439418"
translation="MVSTPENRQTFITSVIKELROYFDFGLDFDWEYFGSRGSPQDK
HLFTVLQVQMRFAFQEAQKINKPRIMVTVAAGISNTQSGYEIPQLSQYLDYIHVM
TYDLHNSPNTGSGPLKYPTDTGSNAYLVNDVYVMYKONGAPAEKLI VGFPTYG
HNFILSNSTGAGTSGAGPAGPVAKESGLWAYEICTFLKNGATOGMDAPQEVY
AYQNVWGYDNIKSFEDIKAWLKNKFGAMVAIDLDDFTGTCNOKKFLISTLK
KALGQASCTAPADIEPIITAPSGNGSGSSGSSGSGSGGSGGFCVAVRANGLYPVAN
NRNPFVHCYNGVTYQQNCQAGLYVFDTSDDCCNWA"
BASE COUNT 327 a 393 c 341 g 308 t
ORIGIN

Alignment Scores:
Pred. No.: 7.17e-156 Length: 1369
Score: 2137.00 Matches: 1369
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 85.28% Indels: 0
DB: 6 Gaps: 0

US-10-004-219b-14 (1-455) x AX405989 (1-1369)
QY 65 LysAsnSerGlnLeuLysThrLeuLeuAlaLeGlyGlyTyrAsnPheGlyThrAlaPro 84
Db 85 CGGAACAGCAGCTGAAACTCTCTCGCCATTTGGAGCTGGAACTTCGGACTGCCCT 144
QY 85 PheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLys 104
Db 145 TTCACCTGCCATGTTCTACTCTGAGAACCGCCAGACTTTCATCACTCACTCACTCAAA 204
QY 105 PheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg 124
Db 205 TTCCTCGCCAGTATGAGTTGACGGGCTGGACTTTGACTGGGAGTACCTGGCTCGT 264
QY 125 GlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAla 144
Db 265 GGAGAGCCCTCTCAGGACAGCATCTCTTCACTGTCTGTGTGAGGAATGCGTGAAGCT 324
QY 145 PheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAla 164
```

```
Db 325 TTTGAGCAGGAGGCAAGCAGATCAACAGCCAGGCTGATGGTCACTGCTGCACTAGCT 384
QY 165 AlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAsp 184
Db 385 GCTGGCATCTCAATATCCAGTCTGGCTATGAGATCCGCCAACTGTACAGTACCTGGAC 444
QY 185 TyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsn 204
Db 445 TACATCCATGTTCATGACCTCAGACCTCCATGGCTCTGGAGGGCTACACTGGAGAGAC 504
QY 205 SerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyr 224
Db 505 AGCCCCCTCTCAAAATACCGACTCAGCCGACCAACCGCTACCTCAATGTGATTAT 564
QY 225 ValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPhePro 244
Db 565 GTCATGAATCTACTGGAGGACAAATGGAGCAGCAGCTGAGAAGCTCATCGTGGATTCCCT 624
QY 245 ThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThr 264
Db 625 ACCTATGGACACAACTTCATCTCTGAGCAACCCCTCCAACTGGAATTGGTCCCCACC 684
QY 265 SerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGlu 284
Db 685 TCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
QY 285 IleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValPro 304
Db 745 ATCTGTACTCTTCTGAAATAATGGAGCCACTCAGGATGGGATGGGCTCAGGAGATGCT 804
QY 305 TyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPheAspIleLys 324
Db 805 TATGCTTATCAGGCAATGTGTGGTGGCTGCTATGCAACATCAAGAGCTTCGATATTAA 864
QY 325 AlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAsp 344
Db 865 GCTCAATGGCTTAAAGCAACAAATAATGGAGCGGCATGCTGCTGGGCCATTCATCTGG 924
QY 345 AspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLys 364
Db 925 GACTTCACTGGGACATTTCTGCAACCAAGGCAAGTTTCCCTAATCTCCACCTCGAAGA 984
QY 365 AlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGluProIleThr 384
Db 985 GCGCTCGGCTCGAGGTGCAAGTTGCAAGCTCCAGCTCAGCCCATTCAGGCAATAACT 1044
QY 385 AlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGlyCysSerGly 404
Db 1045 GCTGCTCCAGTGGGAGCGGGAACGGAGCGGAGTAGCAGCTCTGGAGGCACTCGGGA 1104
QY 405 GlySerGlyPheCysAlaValAlaArgAlaAsnGlyLeuTyrProValAlaAsnArgAsn 424
Db 1105 GGCAGTGGGATCTGTGCTGTGAGCAACAGGCTCTACCCGTGGCAATTAACAGAAAT 1164
QY 425 AlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuVal 444
Db 1165 GCTTCTGCACTGCTGATGAGTCACTACAGCAGCAACTGCCAGCCGCGCTTGTG 1224
QY 445 PheAspThrSerCysAspCysCysAsnTrpAla 455
Db 1225 TTCGACACAGCTGTGATTGCTGCAACTGGCA 1257

RESULT 3
AB025008 1354 bp mRNA linear PRI 25-NOV-1999
LOCUS Homo sapiens TSA1902-L mRNA for novel member of chitinase family,
DEFINITION complete cds.
ACCESSION AB025008
VERSION AB025008.1 GI:6467176
KEYWORDS novel member of chitinase family; TSA1902-L.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

1 (teites)  
 Saito, A., Ozaki, K., Fujiwara, T., Nakamura, Y. and Tanigami, A.  
 Isolation and mapping of a human lung-specific gene, TSA1902,  
 encoding a novel chitinase family member  
 Gene 239 (2), 325-331 (1999)  
 20018184  
 2 (bases 1 to 1354)  
 Submitted (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co.,  
 Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho,  
 Tokushima, Tokushima 771-0192, Japan [E-mail: saito@otsuka.gr.jp,  
 Tel: 81-88-665-2888, Fax: 81-88-637-1035]  
 Location/Qualifiers  
 1. .1354  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1. .1354  
 /genes="TSA1902-L"  
 157. 1263  
 /gene="TSA1902-L"  
 /codon\_start=1  
 /product="novel member of chitinase family"  
 /protein\_id="BAA86980.1"  
 /db\_xref="GI:6467177"  
 /translation="MUSTPENQTFITTSIKELRVEFDGLDFDWEYPCSGSPQDK  
 HFTVLVQEMREAFEGEAKQINKPLRMVTAANAAGISNTQSEYIEPOLQYLDYHVM  
 TYDLHSGWGYTGNSPLKYPTDGSNAYLVNYYVMYKNGNGAPAEKLIYGFPTYG  
 HNFILSPNTGAGTSAGPAGYKESGWAYEICTFLKNGTQMDAPQEPY  
 KVGQNVGVYDYNVSPDIKAQMLKHNFGGAMVWIDLDFTGTCNQGKFLISTLK  
 ALGLOSACTPAQPIETITAPSGSGNGSGSSGSGSGGFCAGRANGLYPVAN  
 NRNAFWHCYGVYQNCQAGLVDFPSCDCCNWA"

BASE COUNT 311 a 387 c 348 g 308 t

ORIGIN

Alignment Scores:

Pred. No.: 1.73e-154 Length: 1354  
 Score: 2119.00 Matches: 387  
 Percent Similarity: 99.49% Conservative: 1  
 Best Local Similarity: 99.23% Mismatches: 2  
 Query Match: 84.56% Indels: 0  
 DB: 9 Gaps: 0

US-10-004-219B-14 (1-455) x AB025008 (1-1354)

66 AenSerGlnLeuLysThrLeuLeuAlaLeuGlyGlyTyrPasnPheGlyThrAlaProPhe 85  
 91 AACAGCCAGCTGAAACCTCTCTGGCCATTTGGAGGCTGGAACCTTCAGGACTGCCCTTTTC 150  
 86 ThrAlaMetValSerThrProGluAsnArgGlnThrPheThrSerValIleLysPhe 105  
 151 ACTGCCATGTTCTACTCTTGAGAACCCGACACTTTCATCACTCAGTCATCAAAATTC 210  
 106 LeuArgGlnTyrGluPheAspGlyLeuAspPheAspTyrPasnPheGlyThrArgGly 125  
 211 CTGGCCAGTATGAGTTTACGGCGCTGGACTTTGACTGGGAGTACCTTGGCTCTCTGGG 270  
 126 SerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPhe 145  
 271 AGCCCTCTCTCAGGACCAAGCATCTCTTCACTGTCTGGTGCAGGAAATGGCTGAACTTTT 330  
 146 GluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaValAlaAla 165  
 331 GAGCAGAGGCCNAGCAGATCAACAGCCAGGCTGATGCTGCTGCTGAGTAGTGTCT 390  
 166 GlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyr 185  
 391 GGCATCTCCAATATCCAGCTGCTGCTGATGATGATCCCACTGTCCAGTACCTGGACTAC 450  
 186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205  
 451 ATCCATGTATGACCTACGACCTCCATGCTCTCTGGAGGCTTACACTGGAGAGAACAGC 510

206 ProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225  
 511 CCCCTCTCAATATCCGACTGACACCGGCAGCAGCTACTCACTCACTGGATATGTC 570  
 226 MetAsnTyrTyrPheAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThr 245  
 571 ATGAACCTACTGGAAGGACAAATGGAGCACCAGCTGAGAGCTCATGTTGGATTCCTTACC 630  
 246 TyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSer 265  
 631 TATGACACCACTTCTCTGAGCAACCTCCACACTGGAATTTGGTCCCTCCCTCTT 690  
 266 GlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGluIle 285  
 691 GGTGCTGCTCTGCTGGGCCCTATGCCAAGAGTCTGGGATCTGGGCTTACTACAGATC 750  
 286 CysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValProTyr 305  
 751 TGTACTCTTCTGAAAATGGAGCCACTCAGGATGGGATGCCCTCAGGAAGTGCCTTAT 810  
 306 AlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleLysSerPheAspIleLysAla 325  
 811 GCCTATCAGGGCAATGTGTGGTGTGGTATGACACGTCACAGCTTCGATATTAGGCT 870  
 326 GlnTyrLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAlaIleAspLeuAspAsp 345  
 871 CAATGGCTTAAGCACACCAAAATTTGGAGCGCATGGTCTGGGCCATTTGATCTGGATGAC 930  
 346 PheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAla 365  
 931 TTCACTGGCAGCTTCTGCAACCCAGGCGAAGTTTCCCTTAATCTCCACCTGAAGAAGGCC 990  
 366 LeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGluProIleThrAla 385  
 991 CTGTGCTGCAGAGTGCAGTTGTCAGCGCTCAGCTCAGCCATTCAGGCAATACTGCT 1050  
 386 AlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGlyLysSerGlyGly 405  
 1051 GCTCCAGCTGGCAGCGGAGAACCGGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGGAGGC 1110  
 406 SerGlyPheCysAlaValArgAlaGlnGlyLeuTyrProValAlaAsnAsnArgAsnAla 425  
 1111 AGTGGATTCTGTCTGGCAGAGCAACCGGCTCTACCCCGTGGCAAAATACAGAAATGCC 1170  
 426 PheTyrHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPhe 445  
 1171 TTCTGGCAGCTGCTGAATGGAGTACGTCACGACGAGAACTGCCAGCGCGGCTTGTCTTC 1230  
 446 AspThrSerCysAspCysCysAsnTyrAla 455  
 1231 GACACAGCTGTGATTTGCTGCAACTGGGCA 1260

RESULT 4

LOCUS

AF154571

DEFINITION

Mus musculus putative chitinase precursor (YNL) mRNA, partial cds.

AF154571

VERSION

AF154571.1 GI:6934189

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1526)

YNL, A Putative Mouse Chitinase

Unpublished

2 (bases 1 to 1526)

Price, P.A., Harris, S.C. and Williamson, M.K.

Direct Submission

Submitted (26-MAY-1999) Biology, University of California, San

Diego, 9500 Gilman Drive, Mail Code 0368, La Jolla, CA 92093-03687,

## USA

## FEATURES

## source

## Location/Qualifiers

1. .1526  
 /organism="Mus musculus"  
 /strain="ATCC1476668"  
 /db\_xref="ATCC:1476668"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE 1209227"  
 /tissue\_type="skin"  
 <1. .1526  
 /gene="YNL"  
 <1. .1419  
 /gene="YNL"  
 /codon\_start=1  
 /product="putative chitinase precursor"  
 /protein\_id="AAF31644.1"  
 /db\_xref="GI:6934190"  
 /translation="AKLLVLTGLALLLNQLGSAYNLICYFTNWAQYRPLGSLFKPDD  
 INPCLTHLIYAFAGQNNIEITIEWNDVLYKAFNDLKNRNSKLKLLAILGWNFGT  
 APFTMVSTSNQROFTITSVIFKLQYGFGLDLWEPGRSGSPDKHLFTVLVKE  
 MREAPEQBAIESNRLMVAAGVGINIOAGYEIPSLKYLDPIHVMYDPLHGSWE  
 GYTGENSLYKYPTEGTGNAYLVNVDYVMNYKNGGAPAEKLVGVPEYGHFTILNPS  
 DNGIGAPSGDGPAGPYTRQAGFWAYEICTFLRSGATEVWDSEVPYAKANEWLG  
 YDNISFSVKAQMLKNQFPGGAWIWDLDLDFGSCDQGFPLTSLINKALGISTEG  
 CTAPDVPEPVTTPFGSG  
 TYQHCQAGLVFDTSNCNCNP"  
 <1. .60  
 /gene="YNL"  
 61..1416  
 /gene="YNL"  
 /product="putative chitinase"  
 394..417  
 /gene="YNL"  
 /note="putative conserved region; active site"  
 415..417  
 /gene="YNL"  
 /note="putative invariant Glu; active site"

## sig\_peptide

## mat\_peptide

## misc\_feature

## misc\_feature

## BASE COUNT

## ORIGIN

## Alignment Scores:

## Pred. No.:

## Score:

## Percent Similarity:

## Best Local Similarity:

## Query Match:

## DB:

## US-10-004-219B-14 (1-455) x AF154571 (1-1526)

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

## QY

## DB

121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
 421 CCAGGCTCAGCTGGGAGCCCTCCTCAGGACAAGCATCTCTTCACTGTCTCTGGTGAAGGAA 480  
 141 MetArgGluAlaPheGluGlnGlnAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
 481 ATGCGTGAAGCTTTTGGAGCAGGAGCTATTGAGACCAACAGCCAGAGCTGATGGTTACT 540  
 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
 541 GCTGCTGAGCTGGTGGGATTTCCAAATCCAGGCTGCTGCTATGAGATCCCTGACTTTCT 600  
 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
 601 AAGTACCTGGATTTCCATGTCATGATATGATGATGATGATGATGATGATGATGATGATG 660  
 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
 661 ACTGGGAGAAATAGTCTCTTTTACAAATACCTACTAGAGCTGGTAGCAATGCCTACCTC 720  
 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240  
 721 AATGGGATTTATGTCATGAACTATTGGGAAGAACATGAGGCCAGCTGAGAGCTCATTT 780  
 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260  
 781 GTTGGATTTCCAGAGTATGGACACACCTTTCCTCTGAGAAACCCCTCTGATAATGGAAT 840  
 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280  
 841 GGTGCCCTTACCTCTGGTGGTGGCTCTGTCGGCCCTATACCCAGCAGAGCTGGTCTGG 900  
 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300  
 901 GCCTACTATGAGATTGACCTTCTGAGAGTGGAGCCACTGAGGCTTGGATGCTCTCC 960  
 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320  
 961 CAAAGAAGTGCCCTATGCTCTATAAGGCCAACAGAGTGGCTTGGCTATGACAAATATCAAGAGC 1020  
 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340  
 1021 TTCAGTGTAAAGGCTCAGTGGCTTAAGCAGACAAATTTTGGAGGTGCCATGATGCTGGGCC 1080  
 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360  
 1081 ATTGACCTTGATGACTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140  
 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380  
 1141 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCTCCTGACGTGCTTCC 1200  
 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGly 400  
 1201 GAGCCAGTACTACTCTCTCCA-----GGAAGTGGAGTGGGGGTGGAAGCTCCGGA 1251  
 401 GlySerSerGlyGlySerGlyPheCysAlaValAlaGlnAsnGlyLeuTyrProValAla 420  
 1252 GGAAGCTCTGGAGGAGTGGATTTCTGTCGCAACAAAGCAGATGGGCTCTTACCTCTGGGCA 1311  
 421 AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnAsnCysGln 440  
 1312 GATGACAGAAATGCTTTTGGCAGTGCATCATGAAATCACAATACACAGCAGCATGTCTCA 1371  
 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454  
 1372 GCAGGGCTTGTGTTGATACAGCTGTAATGCTGCAACTGG 1413

## RESULT 5

## BC011134

## LOCUS

## DEFINITION

## Mus musculus, similar to eosinophil chemotactic cytokine, clone

## BC011134

## LOCUS

## DEFINITION

## Mus musculus, similar to eosinophil chemotactic cytokine, clone

MGC:18771 IMAGE:4165150, mRNA, complete cds.  
 BC011134  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 house mouse;  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1538)  
 Strausberg, R.  
 Direct Submission  
 Submitted (25-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 24 Row: e Column: 11.  
 Location/Qualifiers  
 i. 1538  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="FVB/N"  
 /clone="MGC:18771 IMAGE:4165150"  
 /tissue\_type="Salivary gland, 10 week old female mouse"  
 /clone\_lib="NCI\_CGAP\_SG2"  
 /lab\_host="DH10B"  
 /notes="Vector: pCMV-SPORT6"  
 323..1420  
 /codon\_start=1  
 /product="Similar to eosinophil chemotactic cytokine"  
 /protein\_id="AAH1134.1"  
 /db\_xref="GI:15029822"  
 /translation="MVSTQNRQTFTTSVIFKRLQYFGDLDLWEYPSGRSPQDK  
 HLTFLVEMREAFEOEASNRPRLMVTAAGVGSNTQAGVEIPELSKYLDIFHVM  
 TYDLHGSWGYTGENSEPLVYKYPETGNSAYLVNDVYVMYKNGAPAEKLIIVGFPEY  
 HTFLNPSDNGIGTSGDGPAGPYTROAGFWAYVEICTFLRSQATEVWDASQEVY  
 AKANRWLGDNIKFSVZAQWLKQNFQGMIMWALDDFTGSCFDQCKFLSTLN  
 KALISTECTAPDVSEPVPTTPPGSGSGSGSGSGSGSGSGGFCADKADGLYPVADRN  
 AFQICINGITYQHQAGLVFDTSCNCCNP"  
 BASE COUNT 393 a 382 c 365 g 398 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,23e-153 Length: 1538  
 Score: 2097.50 Matches: 370  
 Percent Similarity: 90.31% Conservative: 40  
 Best Local Similarity: 81.50% Mismatches: 41  
 Query Match: 83.70% Indels: 3  
 DB: 10 Gaps: 1  
 US-10-004-219b-14 (1-455) x BC011134 (1-1538)  
 QY 1 TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
 Db 62 TACAATCTGATGCTATTTTACCACTACCTGGGCCCACTATCGGCACAGTCTGGGAGCTTC 121  
 QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40

122 AAGCCTGATGACATTAACCCCTCGCTGTGTACTACCTGATATGCTGCTGGGATG 181  
 QY 41 GlnAsnAsnGluIleThrThrIleGluTyrAspValThrLeuTyrGlnAlaPheAsn 60  
 Db 182 CAGAACATGAGATCACCACCATAGATGAATGATGATGATCTCTATATAGATCTTCAAT 241  
 QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80  
 Db 242 GACTTGAAAAACAGGACAGCAAACTGAAACCTCTCTGGCAATTTGGAGGCTGGAATTT 301  
 QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
 Db 302 CGAATCTGCTCTCTTCTACCTACCATGCTGTTTCCACTTCTCAGAACCCGACACCTTCATTACC 361  
 QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
 Db 362 TCAGTCATCAAAATTTCTGCGTCAGTATGGTTGATGGACTGGACCTGGAGCTGGGAATAC 421  
 QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
 Db 422 CCAGGCTCAGTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGTCTCTGGTGAAGGAA 481  
 QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
 Db 482 ATGCGTGAAGCTTTTTCAGCAGGAGGCTATTGAGAGCAACAGGCCACAGACTGATGTTACT 541  
 QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
 Db 542 GCTGCTGATGCTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACTTCT 601  
 QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
 Db 602 AAGTACCTGGATTTTCAATGTCATGACATATGACTCCATGCTCTCTGGGAGGCTAC 661  
 QY 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
 Db 662 ACTGGGGAATAGTCTCTTTTACAAATACCTACTAGACTGGTGGTGAATGGAATTT 721  
 QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240  
 Db 722 AATGTGGATTATGTCATGAACCTATTGGGAAGAACAAATGAGGCCCCAGCTGAGAGTCAAT 781  
 QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260  
 Db 782 GTTGGATTTCCAGAGATATGGACACACCTTCTCTGAGAAACCCCTCTGATAATGGAATTT 841  
 QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280  
 Db 842 GGTGCCCTTACCTCTGTGTGATGGCCCTGTGCGGCCCTATACACAGACAGCTGGTCTGG 901  
 QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300  
 Db 902 GCCTACTATGAGATTGACCTTTCTGAGAGTGGAGCCACTGAGGTCTGGATGCTGCTCC 961  
 QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320  
 Db 962 CAAGAAGTGCCTATGCTATAGGCCCAACGAGTGGCTGGCTATGACAAATACAGAGC 1021  
 QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340  
 Db 1022 TTCACTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCC 1081  
 QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360  
 Db 1082 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGGAATTTCTCTGACTCTCT 1141  
 QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380  
 Db 1142 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTCACAGCTCTCTGACGTGCTTCC 1201  
 QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400

Db 1202 GAGCCAGTACTCTCTCCA-----GGAGTGGGAGTGGGGTGGAGCTCCGGA 1252

Qy 401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420

Db 1253 GGAAGCTCTGGAGGAGTGGATCTGTGGCGACAAGAGAGAGGCTTACCTCTGTGGCA 1312

Qy 421 AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440

Db 1313 GATGACAGAAATGCTTTTGGGAGTGCATCAATGGAATCACATACCAGCAGCATGTGCA 1372

Qy 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTyr 454

Db 1373 GCAGGCTTGTGTTTATACACAGCTGTAATGTGCACTGG 1414

RESULT 6

LOCUS BC034548

DEFINITION Mus musculus, chitinase, acidic, clone MGC:19045 IMAGE:4189080, mRNA, complete cds.

ACCESSION BC034548

VERSION BC034548.1 GI:21961190

KEYWORDS MGC.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1557)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [angbcm.tmc.edu](mailto:angbcm.tmc.edu)

Goonaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 24 Row: h Column: 16

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

Location/Qualifiers

1. .1557

/organism="Mus musculus"

/db\_xref="taxon:10090"

/map="FVB/N"

/clone="MGC:19045 IMAGE:4189080"

/tissue\_type="Salivary gland, 10 week old female mouse"

/clone\_lib="NCI CGAP\_SG2"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

4. .1425

/codon\_start=1

/product="chitinase, acidic"

/protein\_id="AAH34548.1"

/db\_xref="GI:21961191"

/db\_xref="LocusID:81600"

/translation="WAKLLVTLGLALLNAQGSAYNICVPTNWAQYRPGSGPKPD DNPCLCHLIYAFAGMKNNEITTEWVDVLYKAPNDLKNRNSKTLTLLAIGWNFG TAPFTTMVTSQNRQTFTSVIKFLQYGFGLDLWDWEYPGSRGSPQDKHLFTVLVK

## FEATURES

source

## CDS

BASE COUNT 410 a 382 c 366 g 399 t

ORIGIN

Alignment Scores:

Pred. No.: 9,37e-153 Length: 1557

Score: 2097.50 Matches: 370

Percent Similarity: 90.31% Conservative: 40

Best Local Similarity: 81.50% Mismatches: 41

Query Match: 83.70% Indels: 3

DB: 10 Gaps: 1

US-10-004-219B-14 (1-455) x BC034548 (1-1557)

Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe 20

Db 67 TACAATCTGATATGCTATTTACCAACTGGGCCAGATATCGCCAGGCTCGGGAGCTTC 126

Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40

Db 127 AAGCCTGATGACATAACCCCTGCTGTACTACCTGATCTATGCTCTGCGGATG 186

Qy 41 GlnAsnAsnGluIleThrThrIleGlnTyrAsnAspValThrLeuTyrGlnAlaPheAsn 60

Db 187 CAGAACATGATGATCACCACCATAGATGATGATGATGATGATGATGATGATGATGAT 246

Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuAlaIleGlyGlyTyrAsnPhe 80

Db 247 GACTTGAATAACAGGACAGCAACTGAAACCTCTCGGCAATTTGGAGCTTGGAACTTT 306

Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100

Db 307 GGAACCTGCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCCGCGACACTTCATTACC 366

Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120

Db 367 TCAGTCATCAAAATTTCTCGTCAGTATGGTTTGTAGTACCTGACCTGGACTGGGAATAC 426

Qy 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140

Db 427 CCAGGCTCAGCTGGAGAGCCCTCTCAGCAACAGCATCTCTTCACTGTCTGCTGGTGAAGNA 486

Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160

Db 487 ATGCTGAAGCTTTTGGAGCAGGAGCTATTGAGACCAACAGGCCCCAGACTGATGTTACT 546

Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180

Db 547 GCTGCTGTAGCTGTGGGATTTCCAAACATCCAGGCTGCTATGAGATCCCTGAACTTTCT 606

Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200

Db 607 AAGTACCTGGATTTTCATTCATGTCATGATGATGATGATGATGATGATGATGATGATGAT 666

Qy 201 ThrGlyGluAsnSerProLeuTyrIleSerThrProThrAspThrGlySerAsnAlaTyrLeu 220

Db 667 ACTGGGGAGATAGTCTCTCTTACAAATACCTACTGAGACTGGTAGCAATGCCTACCTC 726

Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240

Db 727 AATGTGGATTATGTCATGAACCTATTGGGAAGAACAAATGGAGCCCCAGCTGAGAAGCTCAT 786

Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260

Db 787 GTTGATTTCCAGAGTAGTAGGACACACCTTCATCTCGAAGAACCCCTCTGTGATTAATGGAATT 846

Qy 261 GlyAlaProThrSerGlyValaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280





```
Db 726 AATGTGGATTATGTCATGAACCAATGAGAGCCAGCCAGCTGAGAAGCTCAT 785
Qy 241 ValGlyPheProThrTyrrGlyHisAenPheIleLeuSerAsnProSerAenThrGlyIle 260
Db 786 GTTGGATCCAGATGATGACACACCTTCATCTGAGAACCCCTCTGTAATGAAT 845
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrrAlaLysGluSerGlyIleTrp 280
Db 846 GGTGCCCTTACCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
Qy 281 AlaTyrrGluIleCysThrPheLeuLysAenGlyAlaThrGlnGlyTrpAspAlaPro 300
Db 906 GCTACTATGAGATTGGACCTTCTGAGAAGTGGAGCCACTGAGCTGCGGATGCTCC 965
Qy 301 GlnGluValProTyrrAlaTyrrGlnGlyAenValTrpValGlyTyrrAspAenIleLysSer 320
Db 966 CAGAAGTGCCTTATGCTTATAGGCAACAGTGGCTTGGCTTATGACATATCAGAGC 1025
Qy 321 PheAspIleLysAlaGlnTrpLeuLysHisAenLysPheGlyGlyAlaMetValTrpAla 340
Db 1026 TTTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTGGAGTGCCATGATCTGGGC 1085
Qy 341 IleAspLeuAspAspPheThrGlyThrPheCysAenGlnGlyPheProIleLysSer 360
Db 1086 ATTGACCTTGTGATCTCACTGCTCTTCTGTGATCAGGAGAAATTTCTCTGACTCT 1145
Qy 361 ThrLeuLysLysAlaLeuGlnTrpLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
Db 1146 ACTTTGAACAAGACCTTGGCATTCATCAGAGTTCACAGCTCCTGACGTGGCTTCC 1205
Qy 381 GluProIleThrAlaAlaProSerGlySerGlySerGlySerGlySerSerSerGly 400
Db 1206 GAGCCAGTACTCTCTCA-----GGAAGTGGAGTGGGGTGGAGCTCGGA 1256
Qy 401 GlySerSerGlyCysGlyPheCysAlaValArgAlaAenGlyLeuTyrrProValAla 420
Db 1257 GGAAGCTCTGGAGGAGTGGATCTGTGCGCAAAAGCAGATGGCTCTACCTGTGGCA 1316
Qy 421 AenAenAtqAenAlaPheTrpHisCysValAenGlyValTyrrGlnGlnAenCysGln 440
Db 1317 GATGCAAGAAATGCTTTTGGCAGTGATCAATGGAATCACAATACAGCAGCATTTGTCAA 1376
Qy 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
Db 1377 GCAGGCTTTTGTGATACCACTGATTAATGCTGCACTGG 1418

RESULT 8
AB051629 1529 bp mRNA linear MAM 07-NOV-2001
LOCUS Bos taurus cbp b04 mRNA for chitin binding protein b04, complete
DEFINITION cds.
ACCESSION AB051629
VERSION AB051629.1 GI:16754894
KEYWORDS Bos taurus liver cDNA to mRNA.
ORGANISM Bos taurus
REFERENCE 1 Suzuki M., Morimatsu M., Yamashita T., Iwanaga T. and Syuto B.
AUTHORS A novel serum chitinase that is expressed in bovine liver.
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
MEDLINE Bovidae; Bovinae; Bos.
21475601
REFERENCE 2 (bases 1 to 1529)
AUTHORS Suzuki M., Morimatsu M. and Syuto B.
TITLE Direct Submission
JOURNAL Submitted (23-NOV-2000) Maseko Suzuki, Iwate University, Department
of Veterinary Medicine, Faculty of Agriculture; Ueda 3-18-8,
Morioka, Iwate 020-8550, Japan (E-mail: u9998001@iwate-u.ac.jp,
Tel: 81-19-621-6212)
FEATURES Location/Qualifiers
```

```
source 1..1529
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="liver"
gene 1..1529
/feature="cbp b04"
CDS 13..1431
/gene="cbp b04"
/codon_start=1
/product="chitin binding protein b04"
/protein_id="BAB71805.1"
/db_xref="GI:16754895"
/translation="MAKIFITGLAFLNLAQLGSAYLVYPSNWAQVRPGLGSPKPD
NIDPCLPHLYAFAGMNSSETTTTMDNDVALYSSFDLKKNSQLKILATGGWNFG
TAPPTAMVAPENRTFTSSVTKPLHGYDFGLDFDWEVPGPRGSPSODKHLFTVLVQ
ETRAFEQEAOKTNKPRLLVTAAGAAGIQLQAGVEIPLSQYLDPIHMTYDFHGSW
EGYTGENSEPLKYPTDTSNTYLVNEYAMNWKNGAPAEKLIIGFPIYGHNFILRDA
SNNGIGTSGAGPAGPTVREAGWAYEIECAFLKDGATEAWDDDSQNPYATYGTETWV
GYDYNVSPRIKAQWLKENFGGMVWALDDFTGTCFCNQKFPPLINTLKDALGKSA
TCNASTQSSEPNSSPGNESGSGNKSSESSEGRGYCAGKADGLYPVADNRNFAFNVCNGI
TVKQNCILGLVDFDTSCHCNNA"
polya_signal 1508..1513
/gene="cbp b04"
BASE COUNT 391 a 405 c 365 g 368 t
ORIGIN
Alignment Scores:
Pred. No.: 2,996-148 Length: 1529
Score: 2039.00 Matches: 365
Percent Similarity: 86.59% Conservative: 29
Best Local Similarity: 80.22% Mismatches: 57
Query Match: 81.36% Indels: 4
DB: 4 Gaps: 1
US-10-004-219B-14 (1-455) x AB051629 (1-1529)
Qy 1 TyrGlnLeuThrCysTyrrPheThrAsnTrpAlaGlnTyrrArgProGlyLeuGlyArgPhe 20
Db 76 TACCAGCTGGTATGCTACTTCTTAACCTGGGCCAGTACCGTCCAGGCCCTGGGAAGCTTC 135
Qy 21 MetProAspAenIleAspProCysLeuCysThrHisLeuIleTyrrAlaPheAlaGlyArg 40
Db 136 AAGCCTGCAACATCGACCCCTCGCTCGCATCACTGATCTATGCTTCCTCGGATG 195
Qy 41 GlnAenAenGluIleThrThrIleGluTrpAenAspValThrLeuTyrrGlnAlaPheAen 60
Db 196 AGTAACAGTGAGATCACCACCATTTGAATGGAATGATGGCTCTCTATAGTTCTTTCAT 255
Qy 61 GlyLeuLysAenLysAenSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAenPhe 80
Db 256 GACCTGAAAAAAGAAAGACAGCCAGCTGAAAAATTCCTCTGGCCATTTGGAGCTGGAATTT 315
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAenArgGlnThrPheIleThr 100
Db 316 GGAACCTGCCCTTTTCACTGCTATGTTGCTCCATCGGAGAACCGCAAGACTTTTCAATCT 375
Qy 101 SerValIleLysPheLeuArgGlnTyrrGluPheAspGlyLeuAspPheAspTrpGluTyrr 120
Db 376 TAGTCATCAAAATTCCTACACCATGATGGATTTGACGGCTGGATTTTGAATTTGGAGATAC 435
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 436 CTTGCTTTTGGTGGAGGCCCTTCTCAGGACCAAGCATCTCTTCACTGCTCTTAGTGAGGAA 495
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAenLysProArgLeuMetValThr 160
Db 496 ACACGTGAAGCTTTTGAACAGAGAGCAAGCAACCAACCAAGCCAGGCTGCTGGTCCAC 555
Qy 161 AlaAlaValAlaAlaGlyLysSerAenIleGlnSerGlyTyrrGluIleProGlnLeuSer 180
Db 556 GCTGAGTAGCTGCTGGGCTATCTTAATATCCAGGCTGGCTATGAGATTTCCCGAGCTATCC 615
Qy 181 GlnTyrrLeuAspTyrrIleHisValMetThrTyrrAspLeuHisGlySerTrpGluGlyTyrr 200
```

Db 616 CAGTACCTGGACTTCATCCATGTCATGACCTATGACTTCATCGCTCTCTGGGAAGGCTAC 675  
Qy 201 ThrGlyGluAsnSerProLeuTyrGlyTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
Db 676 ACAGGAGAACAGACGCCCTCTCTACAATACCCAACTGACACTGGCAGCAATACTTACCTTC 735  
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyValAlaProAlaGluLysLeuIle 240  
Db 736 ATGTGGAATACGCCATGAACTACTCGAAGAAAATGGGCCCCCGCTGAGAAGCTCATC 795  
Qy 241 ValGlyPheProThrTyrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyIle 260  
Db 796 ATTGATTCCTCGCTACGACACAAATTCATCTAAGAGACGCTTCCAAACAATGGTATC 855  
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTTP 280  
Db 856 GGTGCCCCCATTCTGGAGTGTCTGCTGCGGGCCCTACACAGGAGGCTGGGTCTGG 915  
Qy 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300  
Db 916 GCCTACTATGAGATCTGTGCTTCTTGAAAGATGGAGCCACTGAGGCATGGGATGACTCT 975  
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320  
Db 976 CAGATGTGCTCTATGCTACAGGCGCACTGAGTGGGTGGTTATGACAATGTTAACAGC 1035  
Qy 321 PheAspIleValAlaGlnTrpLeuLysHisAsnLysPheGlyValMetValTrpAla 340  
Db 1036 TTCAGAAATCAAGGCTCAGTGGCTTGAAGGAGAACTTTGGAGGTGCCATGGTCTGGGCC 1095  
Qy 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360  
Db 1096 ATTGACCTGGATGATTTTCAGGCACTTTCAGCAACTTTCGCAACAGGCGCAAGTTCCTCCCTGATAAT 1155  
Qy 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380  
Db 1156 ACCTTGAAGGATCTCTGCTGCTGAAGAGTGAACCTGTAAATGCCCTCCAGCAGTCCAGT 1215  
Qy 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400  
Db 1216 GAGCCCC-----AACAGCAGTCTCGGAACGAAAGTGGGAGTGGAAATAGAGC 1263  
Qy 401 GlySerSerGlyGlySerGlyPheCysAlaValAlaArgAlaAsnGlyLeuTyrProValAla 420  
Db 1264 AGCAGCTCCGAGGTAGAGATCTGTGCTGGCAAGCTGATGCTCTACCCAGTGGCG 1323  
Qy 421 AsnAsnArgAsnAlaPheThrPheCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440  
Db 1324 GATAACAGAAATGCCCTTCTGGAATCGTAAATGGAATCACAATACAAACAGAAATTCCTG 1383  
Qy 441 AlaGlyLeuValPheAspThrSerCysAspCysAsnTrpAla 455  
Db 1384 ACTGGGCTGTCTTTGACACCACTGTCTACTGTCTGCAACTGGGCA 1428  
RESULT 9  
LOCUS AB025009  
DEFINITION Homo sapiens TSA1902-S mRNA for novel member of chitinase family,  
complete cds.  
ACCESSION AB025009  
VERSION AB025009.1 GI:6467178  
KEYWORDS novel member of chitinase family; TSA1902-S.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (sites)  
REFERENCE Saito, A., Ozaki, K., Fujiwara, T., Nakamura, Y. and Tanigami, A.  
TITLE Isolation and mapping of a human lung-specific gene, TSA1902,  
encoding a novel chitinase family member  
JOURNAL Gene 239 (2), 325-331 (1999)  
MEDLINE 20018184

REFERENCE 2 (bases 1 to 1188)  
AUTHORS Saito, A., Ozaki, K., Fujiwara, T., Takahashi, E. and Tanigami, A.  
TITLE Direct Submission  
JOURNAL Submitted (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co.,  
Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho,  
Tokushima, Tokushima 771-0192, Japan (E-mail: saito@otsuka.gr.jp,  
Tel: 81-88-665-2888, Fax: 81-88-637-1035)  
FEATURES  
source  
1..1188  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
gene  
1..1188  
/gene="TSA1902-S"  
CDS  
150..1097  
/gene="TSA1902-S"  
/codon\_start=1  
/product="novel member of chitinase family"  
/protein\_id="BAA86981.1"  
/db\_xref="GI:6467178"  
/translation="MREAFQEAKQINPRLMTVAAGAISNIOGVEIPLSOYLD  
YIHWNTYDLSGSGEYTGENTGSPLYKYPTDTSNAYLVNVDYVWYKNGAPAEKLI  
PFTYGNFILSNPTGIGTSGAGPAGPAGKESGIWAYVEICTFLKNGATQGWDA  
QEVPAVQGNVWGVNDVKSFDIKQWLKHNKFGAMVWAILDLDFTCTFCNQKPEL  
ISTLKALGLOSASCTAPAQPIETITAAPSGSGSGSSGSGSGSGSGSGSGSG  
YEVANNRNFHVCVNGVTYQNCQAGLVFDTSDCCNWA"  
BASE COUNT 280 a 337 c 309 g 262 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,61e-127 Length: 1188  
Score: 1768.50 Matches: 331  
Percent Similarity: 85.13% Conservative: 1  
Best Local Similarity: 84.87% Mismatches: 3  
Query Match: 70.57% Indels: 56  
DB: 9 Gaps: 1  
US-10-004-219B-14 (1-455) x AB025009 (1-1188)  
Qy 66 AsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPhe 85  
Db 91 AACAGCCAGCTGAACACTCTCTGCGCCATTGGAGCTGGAACCTTCAGGACTGC-CCC--- 146  
Qy 86 ThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLysPhe 105  
Db 146 ----- 146  
Qy 106 LeuArgGlnTyrGluPheAspGlyLeuAspPheAspTyrGluTyrProGlySerArgGly 125  
Db 146 ----- 146  
Qy 126 SerProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPhe 145  
Db 147 -----GAAATGGCTGAAGCTTTT 164  
Qy 146 GluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaIle 165  
Db 165 GAGCAGGAGCGCCAAAGCAGATCAACAAAGCCAGGCTGATGCTCAGTCTGCTGCTAGTGTCT 224  
Qy 166 GlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyr 185  
Db 225 GGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTCACTGACTGACTGACTAC 284  
Qy 186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205  
Db 285 ATCCATGTCATGACCTACGACCTCCATGGCTCTCTGGAGGCTACACTGGAGAGAACAGC 344  
Qy 206 ProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225  
Db 345 CCCCTCTCAAAATACCGCACTGACACCGCGCAACAGCGCTACCTCAATGTGATTATGTC 404  
Qy 226 MetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThr 245  
Db 405 ATGAACTACTGGAGAGACAAATGGAGACACAGCTGAGAGCTCATCTGTGGATTCCCTACC 464

QY 246 TyrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyLeuValAlaProThrSer 265  
 DB 465 TATGACACCACTTCTCTGAGCAACCTTCAACACTGGATTGGTGGCCCACTCT 524  
 QY 266 GlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyLeuTyrAlaTyrGluLeu 285  
 DB 525 GGTGCTGGTCTCTGCTGGGCCCTATGCCAAGAGTCTGGGATCTGGGCTTACTACGAGATC 584  
 QY 286 CysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValProTyr 305  
 DB 585 TGTACTCTTCTGAAATAATGGACCACTAGGATGGGATGGCCCTCAGGAAGTGGCTTAT 644  
 QY 306 AlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnLeuLysSerPheAspIleLysAla 325  
 DB 645 GCTATCAGGCAATGTGTGGGTGGCTATGACACGTCACAGCTTCGATATTAGGCT 704  
 QY 326 GlnTyrPheLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAlaIleAspLeuAsp 345  
 DB 705 CAATGGCTTAAGCACAAACAAATTTGGAGGCGCATGGTCTGGGCCATTGATCTGGATGAC 764  
 QY 346 PheThrGlyThrPheCysAsnGlnGlyLysPheProLeuLysSerThrLeuLysLysAla 365  
 DB 765 TTCCTAGGCACTTCTTGCAACAGGCAAGTTTCCCTTAATCTCCACCTCAGGAAGGCC 824  
 QY 366 LeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGluProIleThrAla 385  
 DB 825 CTTGGCTGCAGTGCAAGTTGCACGGCTCAGCTCAGCCATTGAGCCATTAAGTCT 884  
 QY 386 AlaProSerGlySerGlyAsnGlySerGlySerSerSerGlyGlySerGlyGly 405  
 DB 885 GCTCCAGTGGCAGCGGAAACGGAGCGGAGTAGCAGCTCTGGAGGCACTCGGAGGC 944  
 QY 406 SerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAlaAsnArgAla 425  
 DB 945 AGTGGATTCTGTGCTGGCAGACCAACGCGCTTACCCTGGCAATTAACAGAATGCC 1004  
 QY 426 PheTyrPheCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPhe 445  
 DB 1005 TTCCTGGCACTGGTGAATGAGTCACTACGACAGCAAGTCCAGCGCGGCTTGTCTTC 1064  
 QY 446 AspThrSerCysAspCysCysAsnTyrAla 455  
 DB 1065 GACACCACTGTGATGTCTGCAACTGGGCA 1094

RESULT 10  
 AK098814  
 LOCUS  
 DEFINITION Homo sapiens cDNA FLJ25948 fis, clone STM04207, highly similar to Homo sapiens acidic mammalian chitinase precursor.  
 ACCESSION AK098814  
 VERSION AK098814.1 GI:21758941  
 KEYWORDS oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens stomach mucosa cDNA to mRNA, clone lib:STM  
 clone:STM04207.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Nodinomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and Sugano, S.  
 TITLE NEDO human cDNA sequencing project  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 4250)  
 AUTHORS Sugano, S. and Suzuki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome

Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
 (E-mail:cdna@ims.u-tokyo.ac.jp, Tel.81-3-5449-5286,  
 Fax:81-3-5449-5416)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction and 5'-end one pass sequencing: Institute of Medical  
 Science, University of Tokyo, Laboratory of Genome Structure, Human  
 Genome Center; 3'-end one pass sequencing: RAB; clone selection for  
 full insert sequencing: RAB and Helix Research Institute.

# FEATURES

source  
 1..4250  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="STM04207"  
 /tissue type="stomach mucosa"  
 /clone\_lib="STM"  
 /note="cloning vector: pME18SFL3"  
 BASE COUNT 1262 a 868 c 966 g 1154 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8,69e-125 Length: 4250  
 Score: 1742.00 Matches: 318  
 Percent Similarity: 98.47% Conservative: 3  
 Best Local Similarity: 97.55% Mismatches: 5  
 Query Match: 69.51% Indels: 0  
 DB: 9 Gaps: 0

US-10-004-219B-14 (1-455) x AK098814 (1-4250)

QY 130 AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149  
 DB 3174 GATGTCCTATTATGCTTATTTCTTATTTCTGAGGAATGCGTGAAGCTTTTGAGCAGGAGGCC 3233  
 QY 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169  
 DB 3234 AAGCAGATCAACCAAGCCAGGCTGATGGTCACTGCTGCAGTAGCTGCTGGCATCTCCAAT 3293  
 QY 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet 189  
 DB 3294 ATCCAGTCTGGCTATGAGATCCCCCACTGTCCAGTAGCTACATCATCATGTCATGTCATG 3353  
 QY 190 ThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209  
 DB 3354 ACCTACGACCTCCATGGCTCTCTGGAGGGCTACACTGGAGAGAACAGCCCTCTACAA 3413  
 QY 210 TyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTyr 229  
 DB 3414 TACCCGACTGACACCGGCGCAACCGCTACCTCAATGTGGATTATGTCATGAACACTACTGG 3473  
 QY 230 LysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHisAsn 249  
 DB 3474 AAGGCAATGGAGCACCAGCTGAGAGGCTCATCTGTGGATTCCCTACCTATGGACACAAC 3533  
 QY 250 PheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGlyPro 269  
 DB 3534 TTCATCTCTGAGCAACCCCTCCAACTGGATTGGTGGCCCCCACCCTCTGGTCTGGTCTCT 3593  
 QY 270 AlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGluIleCysThrPheLeu 289  
 DB 3594 GCTGGGCCCTATGCCAAGGAGTCTGGGCTCTGGCTTACTAGGAGATCTGTACCTTCTCTG 3653  
 QY 290 LysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValProTyrAlaTyrGlnGly 309  
 DB 3654 AAAAAATGGAGCCACTCAGGAGTGGATGCCCTCAGGAAGTGCCTTATGCCCTATCAGGGC 3713  
 QY 310 AsnValTyrValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTyrLeuLys 329  
 DB 3714 ATGTGTGGTGGCTATGACAACTCAAGAGCTTCGATATTAAAGCTCAATGGCTTAAG 3773  
 QY 330 HisAsnLysPheGlyGlyAlaMetValTyrAlaIleAspLeuAspPheThrGlyThr 349

```
Db 3774 CACAACAAATTTGGAGCGCCGATGCTCTGGGCCATTGATCTGGATGACTTCACTGGCACT 3833
Qy 350 PheCysAsnGlnGlySerPheProLeuIleSerThrLeuIleValAlaLeuGlyLeuGln 369
Db 3834 TTCTGCAACAGGCGCAAGTTTCCCTTAATCTCCACCTGAAGAAGCCCTCGCCCTGGCAG 3893
Qy 370 SerAlaSerCysThrAlaProAlaGlnProIleGluProIleThrAlaAlaProSerGly 389
Db 3894 AGTGCAAGTTGCACGGCTCCAGCTCAGCCATTGAGCCATAACTGCTGCTCCAGTGCGC 3953
Qy 390 SerGlyAsnGlySerGlySerSerSerGlyGlySerSerGlyGlySerGlyPheCys 409
Db 3954 AGCGGGAACGGAGCGGAGTAGCAGCTCTGGAGGCGAGCTCGGGAGCGCAGTGTCTGT 4013
Qy 410 AlaValArgAlaAsnGlyLeuTyProValAlaAsnAsnArgAsnAlaPheTrpHisCys 429
Db 4014 GCTGTGAGAGCAACGGCTCTACCCCGTGGGCAATTAACAGAAATGCCCTTCTGGCACTGC 4073
Qy 430 ValAsnGlyValThrTyGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCys 449
Db 4074 GTGAATGGAGTCACGTACACGACGAGACTGCCAGGCGGGCTTGTCTTCGACACCAGCTGT 4133
Qy 450 AspCysCysAsnTrpAla 455
Db 4134 GATTGCTGCAACTGGGCA 4151.

RESULT 11
M94584
LOCUS M94584 1525 bp mRNA linear ROD 15-JUN-2001
DEFINITION Mus musculus secretory protein precursor (Yml) mRNA, complete cds.
ACCESSION M94584
VERSION M94584.2 GI:11140876
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1525)
AUTHORS Chang, N.-C.A., Liu, C.-H. and Chang, A.C.
TITLE Molecular characterization of a secretory protein (YM-1)
transiently expressed by activated murine peritoneal macrophages
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1525)
AUTHORS Chang, N.C., Hung, S.I., Hwa, K.Y., Kato, I., Chen, J.E., Liu, C.H. and
Chang, A.C.
TITLE A macrophage protein, Yml, transiently expressed during
inflammation is a novel mammalian lectin
JOURNAL J. Biol. Chem. 276 (20), 17497-17506 (2001)
MEDLINE 21264517
PUBMED 11297523
REFERENCE 3 (bases 1 to 1525)
AUTHORS Chang, N.-C.A., Liu, C.-H. and Chang, A.C.
TITLE Direct Submission
JOURNAL Submitted (27-APR-1993) Microbiology and Immunology, National
Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
112, R.O.C.
REFERENCE 4 (bases 1 to 1525)
AUTHORS Chang, N.-C.A., Liu, C.-H. and Chang, A.C.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2000) Microbiology and Immunology, National
Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
112, R.O.C.
REMARK Sequence update by submitter
COMMENT On Nov 13, 2000 this sequence version replaced gi:202441.
FEATURES
source
1..1525
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="PEC in lambda gt11 and ZAPII"
/dev_stage="adult"
1..1525
/gene="Yml"
13..1209
CDS
```

```
/gene="Yml"
/codon_start=1
/product="secretory protein precursor"
/protein_id="AAB62394.2"
/db_xref="GI:11140877"
/translation="MAKILVTGLAILNLVGLSSYQLMCMVYTSWAKDRPIEGSPKPG
NIDPLCTHLIYAFAGMNEITTYHQDLRYEALNGKDKTELKTLILAIGGWKFG
PAPFSAMVSTPNQRQIFQSVIRFLQYDFGLNDWQYPSGRSPDKHLFSVLVK
EMRKAFBESVEKDIPLLSTGAGIIDVIKSGYKIPELSQSLDYIOVMTYDLHDPK
DGYTGNSPLKSPVDIGKSADLNVDSIISYMKDHGAASEKLIYGFPAVGHFTILSDP
SKTGICAPTISTGPPCKYTDSSGLLAYEVECTFLNEGATEWDAPEQVPAAYQGNWV
GYDNVRSPLKRAQWLKNNLGGAVVWPLMDMDFSGSFCHQRHFFPLTSLKGLDNLHSA
SCKGPY"
sig_peptide 13..75
mat_peptide 76..1206
polyA_signal 1487..1492
polyA_site 1507
BASE COUNT 434 a 343 c 335 g 413 t
ORIGIN
Alignment Scores:
Pred. No.: 4..1e-100 Length: 1525
Score: 1415.00 Matches: 256
Percent Similarity: 80.85% Conservative: 48
Best Local Similarity: 68.09% Mismatches: 72
Query Match: 56.46% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-14 (1-455) x M94584 (1-1525)
Qy 1 TyrGlnLeuThrCysTyPheThrAsnTrpAlaGlnTyArgProGlyLeuGlyArgPhe 20
Db 76 TACCAGCTGATGCTGCTACTATACCACTGGGTAGGACAGGCCCAATAGAGGGAGTTTC 135
Qy 21 MetProAspAsnIleAspProCysIleuCysThrHisLeuIleTyAlaPheAlaGlyArg 40
Db 136 AAACCTGGTAATATATGACCCCTGCTGCTACTACCTGATCTATGCTCTTCTGGAATG 195
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyxGlnAlaPheAsn 60
Db 196 CAGATAATAGATCACTTACACACATGACGACGACTTGGCGTACTATGAGCAATTGAAT 255
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 256 GGTCTGAAAGACAGACACACTGAGCTAAATAAATCTCTCTGGCCATTGGAGGATGGAAGTTT 315
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnAtgGlnThrPheIleThr 100
Db 316 GGACCTGCCCGCTTCAGTGCCATGCTCTACTCTCAGAACCGTCAGATATTCATTGAG 375
Qy 101 SerValIleLysPheLeuArgGlnTyGluPheAspGlyLeuAspPheAspTrpGluTy 120
Db 376 TCAGTTATCAGATTCCTTCGTCATATAACTTTGATGGCTCAACCTGAGCAGTAC 435
Qy 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 436 CCTGGGCTTCGAGGAAGCCCTCTTAAGGACAAACATCTCTTCAGTGTTCTGTGTGAAGAA 495
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 496 ATGCGTAAGCTTTTGAGGAAGAAATCTGTGGGAAGAAACATTCAGAGCTGCTACTACT 555
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyGluIleProGlnLeuSer 180
Db 556 TCCACAGGAGCAGGAATCATTGACGCTAATCAAGTCTGGGTACAGATCCCTCAAGTCTCT 615
Qy 181 GlnTyLeuAspTyThrIleHisValMetThrTyArgLeuHisGlySerTrpGluGlyTy 200
Db 181
```



```
QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Db 784 GTGGGATTTCCAGCATATGGGCATACCTTTATCTCGAGTGACCTTCTAAGACTGGAAT 843
QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
Db 844 GGTGCCCTTACAAATTAGTAGTCTGCCGCCACAGGAAAGTACACAGATGAATCAGGACTCCTG 903
QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
Db 904 GCTTACTATGAGGTTTGTACATTTCTGAATGAAGAGGCCACTGAGTCTGGGATGCCCC 963
QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
Db 964 CAGGAAGTACCCTATGCTATCAGGTAATGAGTGGTGGTTTATGACAATGTCAGGAGC 1023
QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyAlaMetValTrpAla 340
Db 1024 TTCAAGTTGAAGGCTCAGTGGCTCAAGGACAAACAATTTAGGAGGTCCCGTGGTCTGCC 1083
QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
Db 1084 CTGACATGATGATCAGTCTGAGTGGTCTTCTGTCCACAGAGACATTTCCCTCTGACATCT 1143
QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaPro 376
Db 1144 ACTTTAAGGAGATCTCAATATACACAGTGCAGTTCGAAGTGGCCCT 1191
RESULT 13
LOCUS AY049765
DEFINITION Mus musculus secreted protein precursor Ym2 mRNA, complete cds.
ACCESSION AY049765
VERSION AY049765.2 GI:22123906
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1209)
AUTHORS Webb,D.C., McKenzie,A.N. and Foster,P.S.
TITLE Expression of the Ym2 lectin-binding protein is dependent on
interleukin (IL)-4 and IL-13 signal transduction: identification of
a novel allergy-associated protein
J. Biol. Chem. 276 (45), 41969-41976 (2001)
JOURNAL 21551268
MEDLINE 11553626
PUBMED 2 (bases 1 to 1209)
AUTHORS Webb,D.C. and Foster,P.S.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Biochemistry and Molecular Biology, John
Curtin School of Medical Research, Australian National University,
Off Mills Rd, Acton, Canberra, ACT 2601, Australia
REFERENCE 3 (bases 1 to 1209)
AUTHORS Webb,D.C. and Foster,P.S.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2002) Biochemistry and Molecular Biology, John
Curtin School of Medical Research, Australian National University,
Off Mills Rd, Acton, Canberra, ACT 2601, Australia
REMARK Sequence update by submitter
COMMENT On Aug 6, 2002 this sequence version replaced gi:15705156.
FEATURES
source
1..1209
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/notes="isolated from the lung lavage from allergic mice"
1..1209
/notes="expression regulated by IL-4 and IL-13 signaling
via the IL-4R alpha subunit"
/codon_start=1
/product="secreted protein precursor Ym2"
```

```
/protein_id="AAL03953.2"
/db_xref="GI:22123907"
/translation="MAKILVLTGAILNLVQLSSQYLMCYTTSWAKDRPTGSGKPKG
NIDPCLCTHLIVAFAGMKNNEITLSEODLRDYDALNLDKDRNLTLLKLLAIGWKFQ
PAPSSMYSVTQNRQTEIKSVIRFLROYNFDGLNLDWOYPGSRGSPDKDHLFSVLVQ
EMRKAFRESEILNHPRIILLSTGAGFDIVIKSGYKIFELSQSLDIYIQWMTYDLHDPK
NGYTGENSPLYSPDYDKSADLNVDSTIITTKDHGAASEKLIIVGFPAHYTFFILSDP
SKNGIGDPTVSAGPPGYKTYNEQLAYFEITCKLNEGATEIFDQEPVFPYALGNBW
GYDNVRSFKLKAQMLKDNLLGAVVWPLDMDDFSGSFCHQGRFPLTTTLTKRDLNVHSA
SCKASYRGEL"
sig_peptide 1..63
BASE COUNT 326 a 273 c 282 g 328 t
ORIGIN
Alignment Scores:
Pred. No.: 7,566-98 Length: 1209
Score: 1384.00 Matches: 250
Percent Similarity: 80.53% Conservative: 52
Best Local Similarity: 66.67% Mismatches: 73
Query Match: 55.23% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-14 (1-455) x AY049765 (1-1209)
QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Db 64 TACCAGCTGATGTGCTACTATACAGCTGGGCTAAGGACAGCCCAACAGAGGGAGTTTC 123
QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Db 124 AAACCTGGTAAATTATGACCCCTGCTGTACTACCTGATCTATGCTCTTCTGGGATG 183
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
Db 184 AGAATATAGATCATTACTTAAGTGAGCAAGACTTGGTGACTATGAACCATTAAT 243
QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 244 GGTCTGAAAGACAGCAACACAGTGCAGTAAACCTCTCTGGCCATTCGAGGATGGAAGTTT 303
QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Db 304 GGACTGCCCTCCCTCAGTTCCAGTCTCTACTCTCAGAACCCGTGACACATTCATTAA 363
QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Db 364 TCAGTTATCAGATCTCTTCGTCATATAACTTTGATGGCTCAACCTGAGCTGGCAGTAC 423
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 424 CCTGGGTCTCGAGGAAGCCCTCTTAAGGACAAACATCTCTTCAGTGTTCGTGGCAGGAA 483
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 484 ATGCTAAAGCTTTTGAGGAAGAATCCACTTTGAACCAACATTCCTCAAGGCTGCTACTCT 543
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Db 544 TCCACAGGAGCTGGATTCATTGACCTTAATCAAGTCTGGGTACAGATCCCTCACTGCT 603
QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 604 CAGTCTCTCGACTATATTTCAGGTTCATGATGATGATGATGATGATGATGATGATGATGAT 663
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 664 ACTGAGAGAAATAGTCCCTCTATAAATCTCATATGATGATGATGATGATGATGATGATGAT 723
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
Db 724 AATGTGGATTCAATTATTACCTACTCTGGAAGGACCATGAGGAGCTTCTGAGAGGCTCAT 783
QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
```

Db 784 GTGGGATTCCAGCATATGGTCTATACCTTTATCTCTGAGTGACCTTCTAAGAATGGAATA 843  
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGlySerGlyIleTrp 280  
Db 844 GTGACCTTACTGTTAGTGTCTGACACACAGGAAAGTACACAATGACAGGACTCTG 903  
Qy 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300  
Db 904 GCTTACTTTGAGATTGTACATTTCTGAATGAAGGAGCCACTGAGATCTTTGATGCCACC 963  
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320  
Db 964 CAGGAAGTACCTATGCTCTATCTGCTGTAATGAGTGGTGGTGTATGACAAATGTCAGGAGC 1023  
Qy 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340  
Db 1024 TTCAGTGTGAAGCTCAGTGGCTCAAGGACAAATATTAGGAGGTGGCTGTCTGGCCC 1083  
Qy 341 IleAspLeuAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360  
Db 1084 CTGGACATGGATGACTTCAGTGGTCTTCTCTCTCACCAGGAGCTTCTCTCTGACAACT 1143  
Qy 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAla 375  
Db 1144 ACTTTAAGAGAGATCTGAATGTACACAGTGCAGAGTTGCAAGGCC 1188  
RESULT 14  
LOCUS AY065557 1209 bp mRNA linear ROD 22-MAY-2002  
DEFINITION Mus musculus putative secretory protein precursor (Ym2) mRNA,  
complete cds.  
ACCESSION AY065557  
VERSION AY065557.1 GI:18086513  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 1209)  
Chang,N.-C.A.  
Tissue-specific Expression of Ym2 Protein  
Unpublished  
2 (bases 1 to 1209)  
Chang,N.-C.A.  
Direct Submission  
TILE  
Submitted (05-DEC-2001) Institute of Microbiology & Immunology,  
National Yang-Ming University, #155 Sec. 2 Li-Nong St., Peitou,  
Taipei 112, Taiwan  
FEATURES  
Location/Qualifiers  
1..1209  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/clone="pGEMT-ez-Ym2.s"  
/tissue\_type="stomach"  
1..1209  
/gene="Ym2"  
1..1209  
/gene="Ym2"  
/codon\_start=1  
/product="putative secretory protein precursor"  
/protein\_id="AAL57751.1"  
/db\_xref="GI:18086514"  
/translation="MAKILVITVGLAILNLVQLSSGYQLMCMYVTSWAKDRPTEGSPKPG  
NIDPCLTLLIYAFAGKNNETIYSEQLRDYEALNGLDKDRNTELEKTLAIGWKFG  
PAPFSMVSPTQNRQTFIKSVIRLQYINFDGLNDWQYPGRPSPKDKILFSLVQ  
EMRKAPEESTLNHPRILLSTGAGFDIVLKSQYKIPELSQSLDIOWMTYDLHDPK  
NGYTGNSPLVSPYDICKSADLNVDLSIITVWKDHGAASEKLIVGFPRAYGHTFLLSDP  
SKNGIGDPTVSGPPGKTNQGLLAIFYEICTFLNEGATEIFDQTEVPYALGNWV  
GTDNVSPFLKAQWLDKNLGVVWPLDMODDFSGSFCHQGRFPPLTTTLTKRDLNVHSA  
SCKASYRGEL"  
326 a 274 c 282 g 327 t  
BASE COUNT

## ORIGIN

Alignment Scores:  
Pred. No.: 1,84e-97 Length: 1209  
Score: 1379.00 Matches: 249  
Percent Similarity: 80.27% Conservative: 52  
Best Local Similarity: 66.40% Mismatches: 74  
Query Match: 55.03% Indels: 0  
DB: 10 Gaps: 0  
US-10-004-219B-14 (1-455) x AY065557 (1-1209)  
Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
Db 64 TACCAGCTGATGTCTACTATACAGCTGGCTGGTAAGGACAGGCCAACAGAGGGAGTTTC 123  
Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40  
Db 124 AAACCTGGTAATATATGACCCCTGCTGTATCTACCTCGATCTATGCTGGGATG 183  
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60  
Db 184 AAGATAATAGATGACACTTACTTAAGTCAGCAGACACTTGGCTGACTATGAAGCATTAAT 243  
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
Db 244 GGTCTGAAAGACAGACACACTGAGCTAAATAAACTCTCTGGCCATTGGAGGATGGAAGTTT 303  
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
Db 304 GGACCTGCCCCGTTTCAGTTCCTCTCTCTCTCAGAACCCGTCAGACATTCATTAAAG 363  
Qy 101 SerValIleLysPheLeuArgGlnTrpGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
Db 364 TCAGTTATCAGATTCCTTCGTCATATAACTTTGATGGCCCTCAACCTGGAGCTGGGAGTAC 423  
Qy 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
Db 424 CCTGGCCCTCGAGGAGGCCCTCTTAAGGACAAACATCTCTTCAGTGTTCCTGGTGAGAA 483  
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
Db 484 ATGCGTAAAGCTTTTGAGGAAGAATCCACTTTTGAACACCATTCCTCAAGGCTGCTACTCT 543  
Qy 161 AlaAlaValAlaAlaGlyLeuSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
Db 544 TCCACAGAGCTGGATTCATTCAGTAATCAAGCTGGGTACAGATCCCTCAAGCTGCT 603  
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
Db 604 CAGTCTCTCGACTATATTCAGTTCATGACATATGATCTCCATGATCTCAAGAATGCTTAC 663  
Qy 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
Db 664 ACTGGAGAAAATAGTCCCTCTATAAATCTCCATATGACATTTGGAAGAGTGTCTGATCTC 723  
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240  
Db 724 AATGGGATTCAAATTTATCTCTGGAAGGACCATGAGGAGCTTCTGAGAGGCTCAT 783  
Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260  
Db 784 GTGGATTTCCAGCATATGGTTCATACCTTTATCTCTGAGTGACCTTCTTAAGAATGGAATA 843  
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGlySerGlyIleTrp 280  
Db 844 GGTGACCTTACTGTTAGTGTCTGACACACAGGAAAGTACACAATGACAGGACTCTCTG 903  
Qy 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300  
Db 904 GCTTACTTTGAGATTGTGACATTTCTGGAATGAAGGAGCCACTGAGATCTTTGATGCCACC 963  
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320







repeat\_region /note="L2 repeat: matches 2595. .2748 of consensus" 37599. .37717

repeat\_region /note="MER94 repeat: matches 1. .119 of consensus" 37777. .38064

repeat\_region /note="L2 repeat: matches 2188. .2502 of consensus" 38106. .38203

repeat\_region /note="MIR repeat: matches 42. .140 of consensus" 39060. .39162

repeat\_region /note="MIR repeat: matches 77. .192 of consensus" 39248. .39475

repeat\_region /note="MIR repeat: matches 17. .261 of consensus" 40505. .40582

repeat\_region /note="L2 repeat: matches 2661. .2740 of consensus" 42696. .42998

repeat\_region /note="Alusg repeat: matches 1. .301 of consensus" 43366. .45787

repeat\_region /note="TRIGER1 repeat: matches 1. .2418 of consensus" 46715. .46908

repeat\_region /note="MIR repeat: matches 1. .206 of consensus" 48333. .49296

repeat\_region /note="LTR5 repeat: matches 1. .968 of consensus" 49597. .49909

repeat\_region /note="Alusg repeat: matches 1. .313 of consensus" 49918. .49957

repeat\_region /note="20 copies 2 mer tt 77% conserved" 49962. .50026

repeat\_region /note="L1P47 repeat: matches 6082. .6145 of consensus" 50027. .50995

repeat\_region /note="LTR5 repeat: matches 21. .969 of consensus" 50996. .51523

repeat\_region /note="L1P47 repeat: matches 5543. .6082 of consensus" 52579. .52919

repeat\_region /note="L2 repeat: matches 2344. .2708 of consensus" 53182. .53681

repeat\_region /note="LTR10D repeat: matches 1. .513 of consensus" 56386. .56623

repeat\_region /note="MIR repeat: matches 2. .215 of consensus" 58241. .58272

repeat\_region /note="16 copies 2 mer aa 84% conserved" 58587. .58978

repeat\_region /note="MSTA repeat: matches 2. .426 of consensus" 59106. .59464

repeat\_region /note="MLTIP repeat: matches 163. .539 of consensus" 59514. .59561

repeat\_region /note="L2 repeat: matches 2576. .2623 of consensus" 59683. .59750

repeat\_region /note="L2 repeat: matches 2635. .2734 of consensus" 60622. .62283

repeat\_region /note="L1M1 repeat: matches -1390. .639 of consensus" 62284. .62605

repeat\_region /note="Alusx repeat: matches 1. .312 of consensus" 62606. .63401

repeat\_region /note="L1M1 repeat: matches 639. .1409 of consensus" 63397. .63569

repeat\_region /note="L1M2 repeat: matches 2559. .2735 of consensus" 63655. .66952

repeat\_region /note="L1M2 repeat: matches 2926. .6308 of consensus" 67388. .67655

repeat\_region /note="L1P83 repeat: matches 5860. .6150 of consensus" 67741. .68108

repeat\_region /note="L1M4 repeat: matches 5483. .5856 of consensus" 68311. .68618

repeat\_region /note="AluY repeat: matches 1. .311 of consensus" 73419. .73525

repeat\_region /note="L2 repeat: matches 2591. .2691 of consensus" 73535. .73690

repeat\_region /note="L2 repeat: matches 2570. .2726 of consensus" 73775. .74035

repeat\_region /note="L2 repeat: matches 2425. .2705 of consensus" 74535. .74690

repeat\_region /note="MIR repeat: matches 46. .205 of consensus"

Pred. No.: 7,47e-94 Length: 91569

Score: 1362.50 Matches: 317

Percent Similarity: 34.12% Conservative: 3

Best Local Similarity: 33.80% Mismatches: 6

Query Match: 54.37% Indels: 613

DB: 9 Gaps: 5

US-10-004-219B-14 (1-455) x AL356387 (1-91569)

Qy 130 AsplysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149

Db 6262 GATGTCCTATTATGCTTATTCTGTAGGAATGCGTAGGCTTTTGAGCAGAGGCC 6321

Qy 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169

Db 6322 AAGCAGATCAACAGCCAGCGCTGATGCTCACTGCTGCAGTAGCTGCTGGCATCTCCAAT 6381

Qy 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGln 181

Db 6382 ATCCAGTCTGGCTATGAGATCCCCCACTGTACAGTGAGTGATGTGCTTATCTTCAA 6441

Qy 181 181

Db 6442 CTCCTGGAGGTGTCCTCGGGCACACTAGACTTGTCTTGGTGTGGCTTATTCAGGG 6501

Qy 181 181

Db 6502 ACCTTGTTTAGGAGACTTTAGNAGTGGTGCTCTGTCCTGCATAGGTCTGGGAATGACC 6561

Qy 182 182

Db 6562 ATCAGAATGATTTTAAATCTCCACCCTCAGGTACCTGGAGTACATCATCTCATG 6621

Qy 190 ThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209

Db 6622 ACCTACGACCTCCATGGCTCTCTGGAGGGCTACACTGGAGAGAACGCCCTCTACAAA 6681

Qy 210 TyrProThrAspThrGlySerAsnAlaTyrLeuAsnVal 222

Db 6682 TACCGAGCTGACACGGCAGCAACGCTTACCTCAATGT-GGTGAGTCCCTGTACAGATGC 6740

Qy 222 222

Db 6801 AAGTTTAGTGGTTTCTCAAATGGGAATTAGGTGGCCATTAATTAATAATATCTCAT 6860

Qy 222 222

Db 6861 TTAATTTACACAGAGACTGGGACACAAAGTAAACATAAATCTGCAAGTGTCTCACITTA 6920

Qy 222 222

Db 6921 AACAACTGTCTCAAGTTGATCGTATTTCATTAACACTACAACCTTACCACAAAATGTTTTGAA 6980

Qy 222 222

Db 6981 AAAAATACTAATCCATCTGGAATTAAACAGCATAGAGTTTTTACCTGTCGGAATAGGGGACT 7040

Qy 222 222

Db 7041 AATATACTAAGTACTGGTCTCTCAGCTGGTGGGCCATGTAACTAACCCACTGACATTG 7100

Qy 223 223

Db 7101 CAGGATTATGTGATGAACACTACTGGAAGGACAATGAGCACCAGCTGAGAAGCTCATCGTT 7160

Qy 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261

Db 7161 GGATTTCCCTACTATGGACAACTTCTCTCTGAGCAACCCCTCCAACTGGAATGGT 7220

262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyLeuTrpAla 281  
Db GCGCCACCTCTGGTGGTCTGCTGGGCGCTATGCCAAGGAGTCTGGGATCTGGGCT 7280  
QY 282 TyrTyrGlu----- 284  
Db 7281 TACTACGAGGTATGTAGATTGGACTGAAAGTGTCTCTGTGAATTCCGTGCACTGTGCCCTT 7340  
QY 284 ----- 284  
Db 7341 AGGCTAGAAATCTGCTGAAATCTTGAATGTTCATCTCTGCTCCTACCTAAATGCTTTAAA 7400  
QY 284 ----- 284  
Db 7401 CACAGCTGCTTAAAGTGTTTAAATAGCAATTCATTACATAATCCAAATGAATAGCATTTA 7460  
QY 284 ----- 284  
Db 7461 ACAATGCATAATATCCCTTCCACACAACAGACTACCTTTTAAATGTAATTTATATTAGCAC 7520  
QY 284 ----- 284  
Db 7521 ACAGGTAATAATAATATATACCTCTCATATTAATCTGTGGGTGCCAAATCATTTGACAA 7580  
QY 284 ----- 284  
Db 7581 GCAACTTGATCCTCTTTACTACAAATAAAACCTGTAACTAGAAATTGACAAACCCCA 7640  
QY 284 ----- 284  
Db 7641 ACTGGAGACATGAAGAAGAAAATTGAAACATGTTTTCTTTAATGGGAGTAACATTT 7700  
QY 285 ----- IleCysThrPheLeuLysAsnGlyAlaThr 294  
Db 7701 AATAGATTTGAATCTTGACTTTTGAAGATCTGTACCTTCTCTGAATAATGAGGCAT 7760  
QY 295 GlnGlyTrpAspAlaProGlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGly 314  
Db 7761 CAGGGATGGATGCCCCCTCAGGAAGTGCCTATATGCCTATCAGGGCAATGTGTGGGTGGC 7820  
QY 315 TyrAspAsnIleLysSerPheAspIleLys----- 324  
Db 7821 TATGACAACATCAAGAGCTTCGATATTAAAGTAAGATCATGCTCCCTTAAATGTGCTGAGT 7880  
QY 325 ----- Ala 325  
Db 7881 CCAGCCCTGAGTCCAGGAACCAAGTATTCCTGTATCTCTTCTTTTAAACAGGCT 7940  
QY 326 GlnTrpLeuLysHisAsnLysPheGlyAlaMetValTrpAlaIleAspLeuAspAsp 345  
Db 7941 CAATGGCTTAAGCACAAACAAATTTGGAGGCGCCATGCTCTGGGCCATTTGATCTGGATGAC 8000  
QY 346 PheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAla 365  
Db 8001 TTCACTGGCACATTTCTGCAACACAGGGCAAGTTTCCCTTAATCTCCACCCTGGAAGAGGCC 8060  
QY 366 LeuGlyLeuGlnSerAlaSer----- 372  
Db 8061 CTCGGCTGCGAGAGTCAAGTAAGTACTGAGGGGGAATGCCCGAGGTGTATTAATACCC 8120  
QY 372 ----- 372  
Db 8121 TTCTCCAACTCAAAAAGCAACCTGTATGTCTTCCACCTCCCTTCCAGGAGATCCTC 8180  
QY 372 ----- 372  
Db 8181 TTTCCACTTCACCTCACCCTCTTGGCCAGATGAGACAGGTTTTTACCCTTGAACGGCC 8240  
QY 372 ----- 372  
Db 8241 ATGTCAAGAACATTTGTCACTTATGGAAGCCAGGCTAACTGGGTGGGTGTTCTAAGTC 8300  
QY 372 ----- 372

Db 8301 TGCTAAGTCTTAAGGGTGTGGAAGAGCTCAAGAAATATATACATCAATATAAGTAT 8360  
QY 372 ----- 372  
Db 8361 CTGCTTCTCTATGAATCCCTTATTCAAAATCTTTCTCAGCCCCACCTGGGGAGATTTAAGA 8420  
QY 372 ----- 372  
Db 8421 AACCTCACCATGCTCACATATCTCTAATCAGTGAATTCAGGGGATGTATCAATCTCTA 8480  
QY 372 ----- 372  
Db 8481 AATTTTTTCCAGATTTTCCCTAGAAATTTATTTGTTTCACTTTATGCCAGATTTCTAGAA 8540  
QY 372 ----- 372  
Db 8541 AAGTGTTTTGCATACATTTGGTGTAGTAAATTTTTTTTGAATGATTTAGATGTATA 8600  
QY 372 ----- 372  
Db 8601 AGGAAAGACCACTTTTTCATAATTAAGCCAGGCTGGCTTGACACAATAGCTTTAT 8660  
QY 372 ----- 372  
Db 8661 TTATCTCTGAATATGGAGTACTTTCTTCTCAGTTGCGAGGATCCAGAGCCCACTGC 8720  
QY 372 ----- 372  
Db 8721 ACCCCACCTCCACACTTCCACTCCACAGTCTCTTCCACTCTAGGTGAAGCTTAGAGCC 8780  
QY 373 ----- CysThrAlaProAla 377  
Db 8781 TCTCCCAAGCCCACTAGTCTGTCTTCTTCTGTGTATGTTTTCAGGTTGCAAGGCTCCAGCT 8840  
QY 378 GlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSer 397  
Db 8841 CAGCCCATTTAGCCCAATAACTGCTCTCCAGTGCAGCGGGAACGGGAGCGGAGTAGC 8900  
QY 398 SerSerGlyGlySerSerSerGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyr 417  
Db 8901 AGCTCTGGAGGCGAGCTCGGGAGGCGAGTGTCTGTGTCTGTCAGAGCCCAACGGCTCTAC 8960  
QY 418 ProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGln 437  
Db 8961 CCGGTGGCAATAACAGAAATGCTTCTGGCACTGCGTGAATGGAGTCACTGACGAGCAG 9020  
QY 438 AsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrpAla 455  
Db 9021 AACTGCCAGGCGGGCTTGTCTTCCACACCAGCTGTGATTGTGCAACTGGGCA 9074

Search completed: June 29, 2003, 23:54:31  
Job time : 3003.5 secs



XX 14-NOV-2000; 2000WO-JP08015.  
XX PR 15-NOV-1999; 99JP-0324467.  
XX PA (TAKE ) TAKEDA CHEM IND LTD.  
XX PI Nakanishi A, Morita S;  
XX WPI; 2001-397791/42.  
XX DR New proteins, peptides and DNA for treatment of bronchial asthma,  
XX PT chronic occlusive lung disease and infectious disease  
XX PS Claim 5; Page 100; 114pp; Japanese.  
XX CC The present invention provides the sequence of a protein which can be  
XX CC used in the treatment and prevention of infectious diseases. Inhibitors  
XX CC of the protein can be used to treat bronchial asthma and chronic  
XX CC occlusive pulmonary disease. The present sequence is an oligonucleotide  
XX CC described in the exemplification of the invention.  
XX SQ Sequence 1368 BP; 324 A; 398 C; 349 G; 297 T; 0 other;

Alignment Scores:  
Pred. No.: 8,966-220 Length: 1368  
Score: 2506.00 Matches: 455  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 22 Gaps: 0

US-10-004-219B-14 (1-455) x AAH42013 (1-1368)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
DB 1 TACCAGCTGACATGCTACTTACCAACTGGGCCAGTACCGGCCAGGCTGGGGCGCTTC 60  
QY 21 MetProAsnAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40  
DB 61 ATGCCTGACAAACATCGACCTCCCTCTGTACCACCTGATCTACGCCCTTGTGGGAGG 120  
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60  
DB 121 CAGAACACAGAGATCACCACCATCGAATGGAATGATGTGACTCTCTACCAAGCTTCAAT 180  
QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
DB 181 GGCTGAAAAAATAAGAACAGCCAGCTGAAACCTCTCTGGCCATTGGAGGCTGGAACTTC 240  
QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
DB 241 GGGACTGCCCTTTTACTGCCATGGTTTCTACTCTCGAGAACCGCCAGACTTTCATCACC 300  
QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
DB 301 TCAGTCATCAATTCCTGCCAGATGATGTTGACGGGCTGGACTTTCAGCTGGAGATAC 360  
QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
DB 361 CTGGCTCTCGTGGGAGCCCTCTCAGGACAGCATCTCTTCACTGCTCTGGTGGAGGAA 420  
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
DB 421 ATGCGTGAAGCTTTGAGCAGAGGCGCAAGCAGATCAACAGCCCAAGGCTGATGTTCACT 480  
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
DB 481 GCTGCAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTCA 540  
QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
DB 541 CAGTACCTGGACTACATCCATGTCATGACCTACGACCTCCATGGCTCTCTGGGAGGCTAC 600

QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
DB 601 ACTGGAGAGAACAGCCCTCTACAAATACCCAGCTGACACCGGAGCAACCCCTACCTC 660  
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240  
DB 661 AATGTGGATTATGTCACTGAACCTACTGGAAGACAATGGAGCACCAGCTGAGAAGCTCATC 720  
QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260  
DB 721 GTTGGATTCCCTTACCTATGGACACAACTTCATCTCTGAGCAACCCCTCCCAACTGGAAT 780  
QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyr 280  
DB 781 GTTGGCCCCACCTCTGGTGTCTGTCTGGCCCTATGCCAAGAGTCTGGGATCTGG 840  
QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300  
DB 841 GCTTACTACGAGATCTGTACCTTCTGAAATATGAGCCACTCAGGATGGATGCCCT 900  
QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320  
DB 901 CAGGAAGTGCCTTATGCCTATCAGGGCAATGTGTGGTGGCTATGACAACTCAAGAGC 960  
QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340  
DB 961 TTCGATATTAAAGCTCAATGGCTTAAGCACACAATAATTTGGAGGCCCATGTCTGGGCC 1020  
QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360  
DB 1021 ATTGATCTGGATGACTTCACTGGCACTTTCTGCAACAGGCAAGTTTCCCCTAATCTCC 1080  
QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380  
DB 1081 ACCCTGAAAGAGGCCCTCGGCTTCAGAGTGCAGTTCAGCTCCAGCTCAGCCCATTT 1140  
QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400  
DB 1141 GAGCCCAATTAATCTGCTCTCCAGTGGCAGCGGGAACGGAGCGGAGTAGCAGCTCTGGA 1200  
QY 401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420  
DB 1201 GGCAGCTCGGAGGCGAGTGGATTCTGTCTGTCTGACAGCAACGGCTCTACCCCGTGGCA 1260  
QY 421 AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440  
DB 1261 AATAACAGAAATGCTTCTGGCAGTGGTGAATGGAGTCACTACCGAGCAAGTGGCCAG 1320  
QY 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrpAla 455  
DB 1321 GCCGGGCTGTCTTCGACACACAGCTGTGATTTGCTGCAACTGGGCA 1365

## RESULT 2

AAH42025  
ID AAH42025 standard; DNA; 1678 BP.  
XX AC AAH42025;  
XX DT 24-AUG-2001 (first entry)  
XX DE Disease treatment related oligonucleotide SEQ ID NO: 16.  
XX KW Disease treatment; infection; chronic occlusive pulmonary disease;  
XX KW bronchial asthma; ds.  
XX OS Homo sapiens.  
XX PN WO200136633-A1.  
XX XX  
XX PD 25-MAY-2001.  
XX PF 14-NOV-2000; 2000WO-JP08015.

XX 15-NOV-1999; 99UP-0324467.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX PI Nakanishi A, Morita S;

XX WPI; 2001-397791/42.

XX New proteins, peptides and DNA for treatment of bronchial asthma,  
PT chronic occlusive lung disease and infectious disease

XX Example 5; Page 107-108; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be  
CC used in the treatment and prevention of infectious diseases. Inhibitors  
CC of the protein can be used to treat bronchial asthma and chronic  
CC occlusive pulmonary disease. The present sequence is an oligonucleotide  
CC described in the exemplification of the invention.

XX Sequence 1678 BP; 397 A; 497 C; 418 G; 366 T; 0 other;

## Alignment Scores:

Pred. No.:	1.19e-219	Length:	1678
Score:	2506.00	Matches:	455
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-10-004-219B-14 (1-455) x AAH42025 (1-1678)

QY	1	TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe	20
DB	206	TTACCAGCTGACATGCTACTTCCAACTAGCGGCCAGTACCGCCAGGCTGGGGCGCTTC	265
QY	21	MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg	40
DB	266	ATGCCTGCAACATCATGACCCCTGCTCTGTACCACCTGATCTAGGCCCTTGTGGGAGG	325
QY	41	GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn	60
DB	326	CAGAACACAGATCATCACCATCGAATCGAATGATGTGACTCTCTACCAAGCTTTCAAT	385
QY	61	GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe	80
DB	386	GGCCTGAAATAAAGAACACCGACGCTGAAACTCTCTGGCCATTGGAGCTGGAACTTC	445
QY	81	GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr	100
DB	446	GGGACTGCCCTTTTCACTGCCATGGTTTCTACTCTGAGAACCGCCAGACTTTTCATCAC	505
QY	101	SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr	120
DB	506	TCAGTCATCAAAATTCCTGGCCAGTATGATTTGACGGGCTGGACTTTGACTGGGAGTAC	565
QY	121	ProGlySerArgGlySerProGlnAspLysPheThrValLeuValGlnGlu	140
DB	566	CCTGGCTCTCGTGGAGCCCTCTCTAGGACAGCATCTCTTCACTGTCTGTGGGAGGAA	625
QY	141	MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr	160
DB	626	ATCGGTGAAGCTTTTGACGAGGAGGCCAAGCAGATCAACAGCCAGGCTGATGGTCACT	685
QY	161	AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer	180
DB	686	GCTGCAGTAGCTGCTGGCATCTCCATATATCCAGTCTGGCTATGAGATCCCCCAACTGTCA	745
QY	181	GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr	200
DB	746	CAGTACTCTGAGTACATCCATGTCATGACCTACGACCTCATGGCTCCCTGGGAGGCTAC	805
QY	201	ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu	220

DB	806	ACTGGAGAGAAACAGCCCTCTTACAAATACCGACTGACCGGAGCAACGCCCTACCTC	865
QY	221	AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle	240
DB	866	AATGGGATTATGTGTAAGTACTTGGAGGAGCAATGGAGCACCAAGCTGAGAGCTCATC	925
QY	241	ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle	260
DB	926	GTGGATTCCCTTACCTATGGACACAACTTCTCTGAGCAACCCCTCCAAACACTGGAAT	985
QY	261	GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp	280
DB	986	GGTGGCCCACTCTGTGGTGGTCTCTGGGCTTATGCCAAGGAGTCTGGGATCTGG	1045
QY	281	AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro	300
DB	1046	GCTTACTACGAGATCTGTACCTTCTGAAATAATGAGGCACCTCAGGGATGGATGCCCT	1105
QY	301	GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer	320
DB	1106	CAGGAAGTGCCTTATGCCTATCAGGGCAATGTGTGGTGGCTATGACACATCAAGAGC	1165
QY	321	PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla	340
DB	1166	TTCCGATATTAAAGGCTCAATGGCTTAAGCACACAAATTTGGAGGCGCCATGCTGGGCC	1225
QY	341	IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer	360
DB	1226	ATTGATCTGGATGACTTCTACCTGGCACTTTCTGCAACACGAGGCAAGTTTCCCTAATCTCC	1285
QY	361	ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle	380
DB	1286	ACCCTGAGAGAGGCCCTCGGCTGCGAGTGCAGTGCAGTGCAGCTCCAGCTCAGCCCAT	1345
QY	381	GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly	400
DB	1346	GAGCCAAATAACTGCTGCTCCAGTGGCAGCGGAAACGGGAGCGGAGTAGCAGCTCTGGA	1405
QY	401	GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla	420
DB	1406	GGCAGCTCGGAGGAGGAGTGGATTTGTGTGTGTCAGAGCAACGGGCTCTTACCCCGTGGA	1465
QY	421	AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln	440
DB	1466	ATAACACAGAAATGCCCTTCTGGCACTGCGTGAATGGAGTCACTACACAGCAACTGCCAG	1525
QY	441	AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrpAla	455
DB	1526	GCCGGGCTTGTCTTCGACACACAGCTGTGATTGTCTGCAACTGGGCA	1570

## RESULT 3

ABN59993

ID ABN59993 standard; cDNA; 1369 BP.

XX AC ABN59993;

XX DT 28-JUN-2002 (first entry)

XX DE Novel human coding sequence SEQ ID NO: 404.

XX KW Human; antianemic; vulnerary; antiinflammatory; immunomodulator;

XX KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;

XX KW neuroprotective; antiparkinsonian; protein therapy; EST;

XX KW expressed sequence tag; gene; ss.

XX OS Homo sapiens.

XX PN WO200222660-A2.

XX XX 21-MAR-2002.

XX PF 10-SEP-2001; 2001WO-US26015.

XX 11-SEP-2000; 2000US-0659671.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
XX Yue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX WPI; 2002-292408/33.  
XX P-PSDB; ABB97580.  
XX  
XX An isolated polynucleotide for treating diseases associated with its  
XX encoded polypeptide such as cancer and multiple sclerosis  
XX  
XX Claim 1; SEQ ID NO 404; 509pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 444  
XX novel human proteins. These were isolated from expressed sequences tags  
XX (ESTs). They can be used to stimulate cell growth, to regulate  
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
XX e.g. in burn treatment, to regulate the immune system e.g. to treat  
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat  
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat  
XX stroke and cancer, to screen for drugs, to treat inflammatory conditions  
XX e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
XX Parkinson's disease. The present sequence is a coding sequence of the  
XX invention.  
XX  
XX Sequence 1369 BP; 327 A; 393 C; 341 G; 308 T; 0 other;  
SQ

Alignment Scores:  
Pred. No.: 5,46e-186 Length: 1369  
Score: 2137.00 Matches: 390  
Percent Similarity: 100.00% Conservatives: 1  
Best Local Similarity: 99.74% Mismatches: 0  
Query Match: 85.28% Indels: 0  
DB: 24 Gaps: 0

US-10-004-219B-14 (1-455) x ABBN59993 (1-1369)

QY 65 LysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrAlaPro 84  
DB 85 CGGACAGCCAGCTGAACTCTCTGGCCATTGGAGGCTGGAACTTCGGGACTGCCCT 144  
QY 85 PheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLys 104  
DB 145 TTCACGCCATGTTCTACTCTGAGAACCCGACACTTTCACCTCAGTCATCAAA 204  
QY 105 PheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg 124  
DB 205 TTCCTGCGCCAGTATGATGTTGACGGCTGGACTTTGACTGGGAGTACCTCGCTCGT 264  
QY 125 GlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAla 144  
DB 265 GGGAGCCCTCTCAGGACAGCATCTCTTCACTGCTCTGGTGAGGAAATGCGTGAAGCT 324  
QY 145 PheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAla 164  
DB 325 TTTGACAGAGGCGCAGCAGATCAACAGCCAGGCTGATGTCAGTCTGTCAGTAGCT 384  
QY 165 AlaGlyLeuSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAsp 184  
DB 385 GCTGGCATCTCCAATATCCAGTCTGGTATGAGATCCCCCACTGTCACAGTACCTGGAC 444  
QY 185 TyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsn 204  
DB 445 TACATCATGTCATGACCTACGACCTCCATGGCTCTCTGGAGGGCTACACTGGAGAAC 504  
QY 205 SerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyr 224  
DB 505 AGCCCTCTCAAAATACCCAGCTGACCCGGCAGCAGCCCTACTCAATGTGGATTAT 564  
QY 225 ValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPhePro 244

DB 565 GTCATGAACACTACTGGAAGGACAAATGAGCACCAGCTGAGAAGCTCATCTTGGATTCCCT 624  
QY 245 ThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThr 264  
DB 625 ACTATGGACACAACCTTCACTCTGAGCAACCCCTCAACACACCTGGAATGGTCCCCCACC 684  
QY 265 SerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGlu 284  
DB 685 TCTGTGCTGCTGCTGCTGGGCCCTATGCCAAGGAGTCTGGGATCTGGGCTTACTACGAG 744  
QY 285 IleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValPro 304  
DB 745 ATCTGTACCTTCTCAAAATGGAGCACTCAGGATGGATGGCCCTCAGGAAGTGCCT 804  
QY 305 TyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPheAspIleLys 324  
DB 805 TATGCTTATCAGGGCAATGTGTGGGTGGCTATGACAACATCAAGAGCTTCGATATTAA 864  
QY 325 AlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAsp 344  
DB 865 GCTCAATGGCTTAAGCACACAAATTTGGAGGCGCATGGTCTGGGCCATTGATCTGGAT 924  
QY 345 AspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLys 364  
DB 925 GACTTCACTGGCACTTCTGTCAACCAAGGCAAGTTTCCCTAATCTCCACCTGAAGAAG 984  
QY 365 AlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGluProIleThr 384  
DB 985 GCCCTCGGCTTCAGAGTGAAGTTGACGGCTCAGCTCAGCCATTGAGCAATAACT 1044  
QY 385 AlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGlyGlySerGly 404  
DB 1045 GCTGTCTCCAGTGGCAGCGGGAACCGGAGCGGAGTAGCAGCTCTGGAGCAGCTCGGGA 1104  
QY 405 GlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAlaAsnAsnArgAsn 424  
DB 1105 GGCAGTGGATTTCTGTGTGTGTCAGAGCAACGGCTCTACCCCGTGGCAATAACAGAAAT 1164  
QY 425 AlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuVal 444  
DB 1165 GCCTTCTGCACTGCTGTAATGGATGAGTACGTAACAGCAGAACTGCCAGCGGGCTTGT 1224  
QY 445 PheAspThrSerCysAspCysCysAsnTrpAla 455  
DB 1225 TTCGACACCACTGTGATTGCTGCACTGGCA 1257

RESULT 4  
AAH42023  
ID AAH42023 standard; DNA; 1469 BP.  
XX  
XX AAH42023;  
AC  
XX  
XX 24-AUG-2001 (first entry)  
DT  
XX  
XX Disease treatment related oligonucleotide SEQ ID NO: 14.  
DE  
XX  
XX Disease treatment; infection; chronic occlusive pulmonary disease;  
KW bronchial asthma; ds.  
XX  
XX Mus sp.  
OS  
XX  
XX WO200136633-A1.  
PN  
XX  
XX 25-MAY-2001.  
PD  
XX  
XX 14-NOV-2000; 2000WO-JP08015.  
PF  
XX  
XX 15-NOV-1999; 99JP-0324467.  
PR  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Nakanishi A, Morita S;  
PI







/product= Mature\_protein

FT XX W09946390-A1.  
 PN XX 16-SEP-1999.  
 PD XX 12-MAR-1999; 99WO-US05343.  
 PF XX 12-MAR-1998; 98US-0039198.  
 PR XX (ICOS-) ICOS CORP.  
 PA XX  
 PI XX Gray PW, Tjoelker LW;  
 DR XX WPI; 1999-551417/46.  
 DR XX P-PSDB; AAY42425.  
 XX  
 PT Novel chitin-binding fragments of human chitinase used to treat fungal  
 PT infections in animals  
 PS Example 1; Page 55-57; 83pp; English.  
 CC This is the nucleotide sequence of an allelic form of the human  
 CC chitinase enzyme, which is capable of degrading Chitin (a linear  
 CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).  
 CC Chitinase fragments can be used to screen for proteins or other  
 CC molecules that specifically bind to the chitin-binding domain of human  
 CC chitinase or that modulate its activity. These compounds are useful for  
 CC immunization, as well as for purifying chitinase, as well as for  
 CC detection and quantification of chitinase. Polynucleotide fragments of  
 CC the invention are useful as a source of probes and primers, and to  
 CC express the proteins recombinantly. The chitinase fragments, when  
 CC conjugated to antifungal compounds, are used to treat animals,  
 CC especially humans, infected with chitin-containing parasites such as  
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,  
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,  
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,  
 CC sporotrichosis, and dermatophytoses.  
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase  
 CC protein for treating infections, especially fungal infections, is  
 CC problematic. In view of the increasing incidents of life-threatening  
 CC fungal infection in e.g. immunocompromised individuals, there exists a  
 CC need for identifying new compounds for treating fungal infection. The  
 CC chitin-binding fragments of the present invention provide this need.  
 XX  
 SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

## Alignment Scores:

Pred. No.: 2,35e-111 Length: 1636  
 Score: 1323.00 Matches: 240  
 Percent Similarity: 67.39% Conservative: 70  
 Best Local Similarity: 52.17% Mismatches: 126  
 Query Match: 52.79% Indels: 24  
 DB: 20 Gaps: 3

US-10-004-219B-14 (1-455) x AAZ21847 (1-1636)

QY 2 GlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21  
 DB 68 AAACGTGGTCTGCTACTTACCAACTGGGCCCCAGTACAGACAGCGGGAGGCTCGCTTCCTG 127  
 QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41  
 DB 128 CCNAGGACTTGGACCCAGCCCTTTGCACCCACCTCATCTAGCCCTTCGTGGCATGCC 187  
 QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61  
 DB 188 AACCCACGACTGAGCACCACCTGAGTGAATGACGAGACTCTTACCGAGGTTCAATGCC 247  
 QY 62 LeuLeuAsnLysAsnSerGlnLeuIleThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
 DB 248 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTTAGCCATCCGAGGCTGGAATTCGCGC 307

QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101  
 DB 308 ACTCAGAAGTTTCACAGATATGCTAGCCACGCGCCCAACACCTGTCAGACCTTTGTCACCTCG 367  
 QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121  
 DB 368 GCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGACTGGAGGTACCCA 427  
 QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141  
 DB 428 GGAAGCAGGGGAGCCCTGCCGTAGACAGAGGCGCTTCACACCTGGTACAGACTTG 487  
 QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161  
 DB 488 GCCAATGCCCTTCCAGCAGGAAGCCGACCTCAGGGAAGGAACGCTTCTTCTGAGTGCA 547  
 QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181  
 DB 548 GCGGTTCCAGCTGGCGACGACCTATGTGGATGCTGATACGAGGTGGACAAATCCGCCAG 607  
 QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201  
 DB 608 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTCAAG 667  
 QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221  
 DB 668 GGACATACAGCCCTCTCAAGAGGCAAGAGAGAGTGGTGCAGCAGCAGCCAGCCTCAAC 727  
 QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241  
 DB 728 GTGGATGCTGCTGTCGAACAGTGTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTT 787  
 QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261  
 DB 788 GCGATGCTTACTACGAGCGCTCTTCACTTGGCTCTCTCATCAGACACAGAGTGGGG 847  
 QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281  
 DB 848 GCCCCAGCAGCAGGCTCGGCACCTCCAGGCCCTTCACCAAGGAAGAGGAGGTGCTGGCC 907  
 QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301  
 DB 908 TACTATGAAGTCTGCTCCTGG-----AAGGGGGGCCACCAACAGAGAAATCCAGGATCAG 961  
 QY 302 GluValProTyrAlaTyrGlnGlyValTrpValGlyTyrAspAsnIleLysSerPhe 321  
 DB 962 AAGTGCCCTACATCTTCGGGACACAGTGGTGGTGGCTTGTATGATGTGGAGAGCTTC 1021  
 QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341  
 DB 1022 AAAACCAAGGTCAAGTATCTGAAGCAGAAAGGAGTGGGCGGGGCCATGGTCTGGGCACTG 1081  
 QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361  
 DB 1082 GACTTAGATGACTTTGCGCGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACG 1141  
 QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374  
 DB 1142 CTACGGCAGGAACTGAGTCTTCCATCTTGGCTTCCAGCACCCAGAGCTTGAAGTTCCA 1201  
 QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394  
 DB 1202 AAACCAAGTTCAGCCCTCTGAACCT----- 1225  
 QY 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414  
 DB 1226 -----GAGCATGGGCCCCAGCCCTGCAAGACACAGTCTTCCAGCGGCAAACTGAT 1276  
 QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434  
 DB 1277 GGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCGCGGGCGGTG 1336  
 QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysAsnTrp 454



342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361  
 1082 GACTTAGATGACTTTCGGCTCTCTCTGCAACAGGCGCCGATACCCCTCTATCCAGACG 1141  
 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374  
 1142 CTACGGCAGGAACCTGAGTCTTCATACCTTCAGGACCCAGAGCTTGAAGTTCCA 1201  
 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394  
 1202 AAACCAAGCTGACGCTCTGAACCT----- 1225  
 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414  
 1226 -----GAGCATGGCCCGCCGCTTGACACAGACGTTCTGCAGGGCAAGCTGAT 1276  
 415 GlyLeuTyProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434  
 1277 GGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTG 1336  
 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454  
 1337 TTCCAGCAAGAGCTGCCCGACAGGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1396

RESULT 8  
 ABL57380  
 ID ABL57380 standard; cDNA; 1636 BP.

ABL57380;  
 12-AUG-2002 (first entry)  
 Human Chitinase cDNA clone MO-218.  
 Chitinase; enzyme; human; fungicide; antifungal; infection;  
 candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;  
 paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
 chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;  
 Pneumocystis; gene; ss.  
 Homo sapiens.

Key Location/Qualifiers  
 CDS 2..1402  
 /tag= a  
 /product= "Chitinase"  
 sig\_peptide 2..64  
 /tag= b  
 mat\_peptide 65..1399  
 /tag= c  
 US6372212-B1.  
 16-APR-2002.  
 16-JUN-1997; 97US-0877599.  
 14-JUN-1996; 96US-0663618.  
 (ICOS-) ICOS CORP.

Gray PW;  
 WPI; 2002-442449/47.  
 P-PSDB; ABB76291.  
 Co-administering chitinase to improve the effectiveness of fungicidal  
 drugs e.g. amphotericin B or itraconazole, useful for treating fungal  
 infections e.g. candidiasis, coccidioidomycosis and blastomycosis -  
 Example 1; Column 23-28; 26pp; English.

The present sequence is that of cDNA clone MO-218 (ATCC 98077) encoding human chitinase (see ABB76291). The clone was isolated from a cDNA library prepared from peripheral blood monocyte-derived macrophages following sequence analysis. Also isolated was clone MO-13B (see ABL57381), which contains a single nucleotide difference in the coding region, changing the encoded amino acid at position 81 of the mature protein from glycine to serine. Northern blots showed highest chitinase gene expression in lung and ovary tissues. Expression in lung is consistent with a protective role against pathogenic organisms that contain chitin. The invention provides human chitinase polynucleotides and polypeptides, and materials and methods for the recombinant production of human chitinase products, which are expected to be useful as products for treating fungal infections or for the development of such products. Human chitinase has a synergistic effect on the actions of other fungicides. It can be administered to improve the antifungal activity of a non-chitinase antifungal agent, especially amphotericin B or itraconazole, in the treatment of a fungal infection such as candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis, dermatophytoses and Pneumocystis infections (all claimed). In particular, the fungal infection involves Candida, Aspergillus and/or Cryptococcus spp., whose growth is not effectively inhibited by contact with human chitinase alone.

Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,35e-111 Length: 1636  
 Score: 1323.00 Matches: 240  
 Percent Similarity: 67.39% Conservative: 70  
 Best Local Similarity: 52.17% Mismatches: 126  
 Query Match: 52.79% Indels: 24  
 DB: 24 Gaps: 3

US-10-004-219B-14 (1-455) x ABL57380 (1-1636)  
 QY 2 GlnLeuThrCysTyThrPheThrAsnTrpAlaGlnTyArgProGlyLeuGlyArgPheMet 21  
 DB 68 AAACCTGGTCTGCTACTTCCCACTGGGCCCATACAGACAGGGGGAGGCTCGCTTCCTG 127  
 QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyAlaPheAlaGlyArgGln 41  
 DB 128 CCCAAGGACTTGGACCCCGAGCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATGCC 187  
 QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyArgGlnAlaPheAsnGly 61  
 DB 188 AACCAACAGCTGAGCACCACCTGAGTGAATGACGAGACTCTTACCAGGAGTTCAATGGC 247  
 QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
 DB 248 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAAATTCGGC 307  
 QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101  
 DB 308 ACTCAGAAGTTTCACAGATATGGTAGCCACGCGCAACACCGTCAGACCTTTGTCAACTCG 367  
 QY 102 ValIleLysPheLeuArgGlnTyThrGluPheAspGlyLeuAspPheAspTrpGluTyPro 121  
 DB 368 GCCATCAGGTTTCTGCGCAAAATACAGCTTTGACGCGCTTGACCTTGAGTGGAGTACCCA 427  
 QY 122 GlySerArgGlySerProProGluAspLysHisLeuPheThrValLeuValGlnGluMet 141  
 DB 428 GGAAGCCAGGGAGGCCCTTCGCGTAGCAAGAGGCGCTTCACACCCCTGGTACAGACTTG 487  
 QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161  
 DB 488 GCCAATGCTTCCAGCAGAGAGCCAGACCTCAGGGAAGGAACGCTTCTTCTGAGTGCA 547  
 QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyThrGluIleProGlnLeuSerGln 181  
 DB 548 GCGGTTCCAGCTGGGCGAGACCTTATGTGATGCTGTGATACGAGGTGACAAATATCGCCAG 607



```
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
D 379 GCCATCAGGTTCTGCGCAATACAGCTTTGACGGCTTGACTTGAGTGGAGTACCCA 438
QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
D 439 GGAAGCCAGGAGGAGCCCTCCCTAGACAGGAGCGCTTCACAAACCTGGTGCAGGACTTG 498
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
D 499 GCCATGCTTTCAGCAGAGAACCCAGACTCAGGAAGAGACGCTTCTTCTGAGTGA 558
QY 162 AlaValAlaAlaGlyLysSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
D 559 GCGGTTCAGCTGGGCAGACCTATGTGATGCTGATACAGGCTTCACAAACCTGGTGCAGGACTTG 618
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
D 619 AACCTGGATTTTGTCAACCTTATGCTTACGACTTCCATGGCTCTTGGGAGAGTCAAG 678
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
D 679 GGACATACAGCCCTCTACAGAGCGAAGAGAGTGTGCGACGACCCAGCCTCAAC 738
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
D 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGAGCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyPheProThrTyrGlyHisAsnPhelIleLeuSerAsnProSerAsnThrGlyIleGly 261
D 799 GGCATGCTTACCTACGGAGCTCTTCAACATGGCTCTTCAACAGAGAGGAGGATGCTGGCC 858
QY 262 AlaProThrSerGlyValAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
D 859 GCCCAGCCAGGAGGTCTGGCACTCAGGCCCTTCAACAGAGAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGln 301
D 919 TACTATGAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAG 972
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
D 973 AAGTGCCCTACATCTTCCGGACACCAACAGTGGTGGCTTTGATGATGTGGAGGCTTC 1032
QY 322 AspileLysAlaGluTrpLeuLysHisAsnLysPheGlyValaMetValTrpAlaIle 341
D 1033 AAACCAAGTCCAGCTATCTGAAGCAGAGAGGAGTGGGGGGGCGCATGCTCTGGGCACTG 1092
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
D 1093 GACTTAGATGACTTTGCGCGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGC 1152
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
D 1153 CTACGCGAGAACTAGTCTTCATACTTGCCTTCCAGGCACCCAGAGCTTGAAGTTCCA 1212
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
D 1213 AAACCAAGTTCAGCCCTCTGAACCT----- 1236
QY 395 GlySerSerSerGlyGlySerGlySerGlyPheCysAlaValArgAlaAsn 414
D 1237 -----GAGCATGGCCCGCCAGCTGTCGACAGACAGCTTCTGCGAGCGCAAGCTGAT 1287
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
D 1288 GGGCTCTATCCCAATCTCTCGGAACGGTCCAGTCTTACAGCTGTGCAGCGGGGGGCTG 1347
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValLeuAspThrSerCysAspCysAsnTrp 454
D 1348 TTCCAGCAAGAGCTGCCCGCAGCGCTGGTGTTCAGCAACTCTCGCAATGTGCTGCACCTGG 1407
```

```
RESULT 10
AAV10436
ID AAV10436 standard; cDNA; 1656 BP.
XX
AC AAV10436;
XX
DT 15-JUN-1998 (first entry)
XX
DE Human chitinase clone MO-13B cDNA.
XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; hybridisation; detection; hybridisation; antifungal;
KW Rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 27..1427
FT /tag= a
FT sig_peptide 27..89
FT mat_peptide 90..1424
FT /tag= c
FT /product= chitinase
FT /note= "from clone MO-13B"
XX
XX WO9747752-A1.
XX
PD 18-DEC-1997.
XX
PF 16-JUN-1997; 97WO-US10460.
XX
PR 14-JUN-1996; 96US-0663618.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW;
XX
DR WPI; 1998-052316/05.
DR P-PSDB; AAW40260.
XX
XX Nucleic acids encoding human chitinase - useful as antifungal
XX agents, especially in combination with other antifungals
XX
PS Claim 9; Page 42-44; 63pp; English.
XX
CC This sequence encodes a novel human chitinase isolated from clone MO-13B.
CC Chitinases are useful for treating or preventing fungal infection and
CC as immunogens for generating antibodies which are used to purify, detect
CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
CC nucleic acid sequence of the chitinase is also useful as a probe to
CC identify and isolate genomic DNA encoding chitinases or similar proteins,
CC or cells expressing them or to generate transgenic ('knockout') rodents.
CC It can also be used in hybridisation assays and to detect genetic
CC alterations in the chitinase gene related to disease. Agents that
CC this protein may be useful in treatment of Gaucher's disease and
CC rheumatoid arthritis, where overexpression of the protein can damage
CC the extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,
CC and thus of their side effects.
XX
SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;
XX
```

```
Alignment Scores:
Pred. No.: 8.46e-111 Length: 1656
Score: 1317.00 Matches: 239
Percent Similarity: 67.17% Conservative: 70
Best Local Similarity: 51.96% Mismatches: 127
Query Match: 52.55% Indels: 24
DB: 19 Gaps: 3
```

US-10-004-219B-14 (1-455) x AAV10436 (1-1656)

```
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
Db : : : : :
93 AAATGCTGCTACTTCAACCACTGGGCCAGGTACAGACAGGGGAGGCTGCTTCTCTG 152

QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
Db : : : : :
153 CCAAGAGACTTGGACCCAGCCCTTTCACCCACTCATCTACGCCCTTCGCTGCGATGACC 212

QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValTrpLeuTyrGlnAlaPheAsnGly 61
Db : : : : :
213 AACACAGCTGAGCACCACTGAGTGGATGACAGACTCTCTACAGAGTTCATATGCG 272

QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
Db : : : : :
273 CTGAAGAAGATGAATCCCAAGCTGAAGCCCTGTAGCCATCGAGGCTGGAATTTTCAGC 332

QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
Db : : : : :
333 ACTCAGAAGTTCACAGATATGGTAGCCAGCCCAACACCGTCAGACCTTTGTCAACTCG 392

QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
Db : : : : :
393 GCCATCAGTTCCTGGCAATATACAGCTTTCAGGCTTGACCTTGACTGGGAGTACCCA 452

QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
Db : : : : :
453 GGAAGCCAGGGAGCCCTGCGGTAGACAGAGGCGCTTCACAACCCCTGGTACAGGACTTG 512

QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
Db : : : : :
513 GCCAATGCCCTTCAGCAGGAAGCCACAGCTCAGGGAAGCAAGCCCTTCTTCGTAGTGA 572

QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
Db : : : : :
573 GCGGTCCAGCTGGCAGACCTATGTGATGTGTGATACAGGTGACAAATCGCCAG 632

QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
Db : : : : :
633 AACCTGGATTTCTCAACCTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTTCAGC 692

QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
Db : : : : :
693 GGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGTGCGACGACGACCCCTCAAC 752

QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
Db : : : : :
753 GTGATGCTGCTGCAACAGTGTGCGAGAGGGACCCCTGCCAGCAAGCTGATCTT 812

QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
Db : : : : :
813 GGCATGCTACCTACGAGCGCTCTTCACACTGGCTCTCCTCCTCATCAGACACAGAGTGGG 872

QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
Db : : : : :
873 GCCCCAGCACAGGGTCTGCACTCCAGGCCCTTCCAAAGGAAGAGGAGTGTGGCC 932

QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
Db : : : : :
933 TACTATGAGTCTGCTCTG-----AAGGGGGCCCAACCAACAGAGATCCAGGATCAG 986

QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
Db : : : : :
987 AAGTGCCCTTACATCTTCGGGCAACACCACTGCGGTGGTGTTCATGATGTGGAGAGCTTC 1046

QY 322 AspIleIleValAlaGlnTrpLeuLysHisAsnLysPheGlyClyAlaMetValTrpAlaIle 341
Db : : : : :
1047 AAACCAAGTCACTATCTGAAGCAGAGAGGGAGTGGCGGGCCCATGTCTGGGCACATG 1106

QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
Db : : : : :
1107 GACTTAGATGACTTTGCGGCTTCTCTCTGCAACAGGCGGATACCCCTCATCCAGAGC 1166

QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
```

```
Db 1167 CTACGCGAAGAACTAGTCTTCATACCTTGCCTTCAGGCACCCAGAGCTTGAAGTTCCA 1226

QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
Db : : : : :
1227 AAACGAGGTCAAGCTCTGAACCT----- 1250

QY 395 GlySerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
Db : : : : :
1251 -----GAGCATGGCCCCAGCTTCTGCAGGGCAAGCTGAT 1301

QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
Db : : : : :
1302 GGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGTG 1361

QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
Db : : : : :
1362 TTCCAGCAAGAGCTGCCGACAGGCTGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1421

RESULT 11
AAZ21848
ID AAZ21848 standard; DNA; 1656 BP.
XX
AC AAZ21848;
XX
XX 10-DEC-1999 (first entry)
XX
DE MO-13B clone of human Chitinase, with noncoding 5'/3' regions.
XX
XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin-binding; allele; vector;
KW truncated protein; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 27..1427
FT /tag= a
FT /product= Human_Chitinase
FT sig_peptide 27..89
FT /tag= b
FT /note= "Signal peptide"
FT mat_peptide 90..1427
FT /tag= c
FT /note= "Mature peptide"
XX
XX WO9946390-A1.
PN
XX
PD 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US05343.
XX
XX 12-MAR-1998; 98US-0039198.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW, Tjoelker LW;
XX
XX WPI: 1999-551417/46.
XX
XX P-PSDB; AAY42426.
XX
XX Novel chitin-binding fragments of human chitinase used to treat fungal
XX infections in animals
XX
XX Example 1; Page 59-62; 83pp; English.
XX
XX This is the nucleotide sequence of an allelic form of the human
XX chitinase enzyme, which is capable of degrading Chitin (a linear
XX homo polymer of beta-1,4-linked N-acetylglucosamine residues).
XX Chitinase fragments can be used to screen for proteins or other
XX molecules that specifically bind to the chitin-binding domain of human
XX chitinase or that modulate its activity. These compounds are useful for
XX immunization, as well as for purifying chitinase, as well as for
```

[illegible][illegible]

DD 387 AAGGTGCCCCACATCTTCCGGGACCAACAGTGGGTGGGCCTTGAATGATGGAGAGCTTC 1048

[illegible]

LD GACTTAAATGACTTTCCTGGCTTCCTCTGCAACCAAGGCCCCCAATACCCCCCTCATCCAGACG 110

Q9      LeuGluGlnAlaLeu-----GluLeuGlnIleGlnAspCysHis 374  
         |||:::      |||  
362 LeuGluGlnAlaLeu-----GluLeuGlnIleGlnAspCysHis 374

DD 1187 C TACGGCAGGAACTGAGTCTTCCATACCTTGCCCTTCAGGCACCCCGAGAGCTTGAAGTTCCTA 1228

QY : 373 ATAFIOATAGINPFOITTEGIUPFOTIETNIAATAFAFOSEGLYSEGLYASNGLYSER 394

DDZ / AAATCCAGGTCAGCCCTCTGAACTT-----1230

[illegible]

DD 1251 -----GAGCATGGCCCCCAGCCCCCTGGACAAAGACACGTTCTGCCAGGGGCAAAAGCTGAT 1301

QY 413 GYUOATATASMSIALYASIMATAFNEIRPHISCYSAATSMYVAIIM 434

DD 130Z 00001C1A1TCCCA1TCCCTC00GAA1CGG1CCAGCT1CTACAGCTGTG1CAGC000000C10 1301

QY  
433 YUGINGJIANCUBGIMATAGYLEUVAIRKLEAPRINIISEICUBASPCUSYBASNIPR 434

.

:

DD 1302 1TCCAGACAAAGCTGCCCCGACAGGCCC1GGG1G1TTCAGCTAC1CC1GCGAAATGCTGCACCC1GG 1421

RESOL 12  
AAD03760

AD03760 Standard; CDNA; 1636 BF.  
XX

AC  
RADUS 780;  
XX

DI T002-NOON-15 (TITLE empty)  
XX  
XX

DE  
XX  
XX  
human cytochrome cDNA from clone pMO-13B.

KW fungal infection: candidiasis: aspergillosis: coccidioidomycosis: human; antifungal; chitinase; immunoglobulin; 19; therapy;

KW chromoblastomycosis: sporotrichosis: mucormycosis: dermatophytoses;  
KW diastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;  
KW chromoblastomycosis: sporotrichosis: mucormycosis: dermatophytoses;

XX  
XX  
XX

[illegible]

EH	Key	Location/Qualifiers
ET	CDS	27..1427

```
ET /cag= a
ET /product= "Human chitinase from clone pMO-13B"
```

```

FT sig_peptide
Z/::89
FT /*tag= b

```

```

FT _perror
FT 30...11223
FT /*tag= c

```

/product= "Human mature chitinase from clone pMO-13B"

WT0200123430-A2.  
05-APR-2001.  
28-SEP-2000; 2000WO-US26960.  
30-SEP-1999; 99US-0409918.  
(ICOS-) ICOS CORP.  
Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;  
WPI: 2001-266141/27.  
P-PSDB; AAE00433.  
Novel chitinase immunoglobulin fusion product, useful for treating fungal infections and reducing the amount of a non-chitinase antifungal agent needed for the treatment  
Claim 2; Page 34-36; 39pp; English.  
The present invention relates to a chitinase immunoglobulin (Ig) fusion product, comprising a human chitinase fused to at least a portion of an immunoglobulin chain. The fusion product is useful for treating fungal infections (mycoses) such as candidiasis, aspergillosis, blastomycosis, coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis, cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses. The fusion protein is useful for reducing the amount of non-chitinase antifungal agent needed to exert an antifungal activity. The fusion protein is also useful for preparing a medicament for the prophylactic or therapeutic treatment of fungal infections. Chitinase immunoglobulin fusion product has unexpectedly improved serum half-life and formulation properties. The present sequence is human chitinase cDNA from clone pMO-13B. Chitinase enzyme degrades chitin which is a homopolymer of beta-(1,4)-linked N-acetylglucosamine residues.  
Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:  
Pred. No.: 8,46e-111 Length: 1656  
Score: 1317.00 Matches: 239  
Percent Similarity: 67.17% Conservative: 70  
Best Local Similarity: 51.96% Mismatches: 127  
Query Match: 52.55% Indels: 24  
DB: 22 Gaps: 3  
US-10-004-219B-14 (1-455) x AAD03760 (1-1656)  
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21  
Db 93 AACTGGTCTGCTACTTACCACTAGTGGCCAGTACAGACAGGGAGGCTGCTTCCTG 152  
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41  
Db 153 CCCAAGGACTTGGACCCAGCTTTGCACCCACCTCATCTACGCCCTTGGTGGCATGACC 212  
QY 42 AsnAsnGluLeuThrThrleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61  
Db 213 AACCCAGCTGAGCACCACCTAGTGGATGACGAGACTCTCTACAGGAGTTCAATGGC 272  
QY 62 LeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
Db 273 CTGAAGAAGTGAATCCAGCTGTAGAACCTTTAGCCATCGAGGCTGGAATTTTCAGC 332  
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101  
Db 333 ACTCAGAAGTTTACAGATATGGTAGCCAGCCGCAACCGTCAGACCTTTGTCAACTCG 392  
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121  
Db 393 GCCATCAGGTTTCTGGCAAAATACAGCTTTGACGGCTTGACCTTGACCTGGGAGTACCCA 452

QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141  
Db 453 GGAAGCCAGGGAGCCCTGCCGTAGCAAGAGGAGCGCTTCAACAACCTGGTGCAGGACTTG 512  
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161  
Db 513 GCCAATGCCCTTCCAGCAGGAAGCCAGCCTCAGGAAGGAACGCGCTTCTCTGAGTGCA 572  
QY 162 AlaValAlaAlaGlyLeuSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181  
Db 573 GCGGTTCCAGCTGGGCAGACCTATGTGGATGCTGATACGAGGTGACAAATGCCCCAG 632  
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201  
Db 633 AACCTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAAGGTTCAG 692  
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221  
Db 693 GGACATAACAGCCCTCTTACAAGAGAGCAAGAGAGTGGTGCAGCAGCCAGCCTCAAC 752  
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241  
Db 753 GTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGACCCCTGCCAGCAAGTGTATCTT 812  
QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261  
Db 813 GGCATGCCCTTACCTACGGAGCGCTCTTCACTGGCTCTCATCATCAGACACCAAGTGGGG 872  
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281  
Db 873 GCCCCAGCCAGGCTTGGCACTCCAGGCCCTTCCACCAAGGAAGAGGAGTGTGGCC 932  
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301  
Db 933 TACTATGAAGTCTGCTCTCTGG-----AAGGGGGCCACCAACAGAGATCCAGATCAG 986  
QY 302 GlnValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321  
Db 987 AAGGTGCCCTACATCTTCCGGGACCAACCACTGGTGGCTTTGATGATGTGGAGACTTC 1046  
QY 322 AsnIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341  
Db 1047 AAAACCAAGGTGAGCTATCTGAAGCAGAGAGGAGCTGGGGCGGGCCATGCTCTGGGCAC 1106  
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361  
Db 1107 GACTTAGTAGACTTTTGGCGGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCCAGACG 1166  
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374  
Db 1167 CTACGGCAGGAAGTGTGCTTCCATCTTCCATCTTCCAGGCACCCAGAGCTTGAAGTTCCA 1226  
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394  
Db 1227 AAACCAAGGTGAGCCCTCTGAACCT----- 1250  
QY 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414  
Db 1251 -----GACATGGGCCCGCCCTGGCAAGACAGCTTCTGCAGGGCGAAGCTGAT 1301  
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434  
Db 1302 GGGCTCTATCCCACTCTCGGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTG 1361  
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454  
Db 1362 TTCCAGCAAGCTGCCCGCAGCGCTGTGTTCAGCAACTCTCTGCAATGCTGCACCTGG 1421  
RESULT 13  
ABL57381  
ID ABL57381 standard; cDNA; 1656 BP.  
XX





QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341  
 Db 1047 AAAACCAAGGTACGCTATCTGAAGCAGAGGGAGTGGGGGCCCATGTGTGGGCACTG 1106  
 QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361  
 Db 1107 GACTTAGTAGACTTTGCGCGCTTCTCTGCAACCGGCGGATACCCCTCATCCAGACG 1166  
 QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374  
 Db 1167 CTACGGCAGGAAGTGAAGTCTTCATCTGCTTCAGGCACCCAGAGCTTGAGTTCCA 1226  
 QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394  
 Db 1227 AAACCAAGGTGAGCCTCTGAACCT----- 1250  
 QY 395 GlySerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414  
 Db 1251 -----GAGCATGGCCCGCAGCCCTGGACACAGACACGTTCTGCCAGGGCAAAGCTGAT 1301  
 QY 415 GlyLeuTyrrProValAlaAlaAsnArgAsnAlaPheTrpHisCysValaAsnGlyValThr 434  
 Db 1302 GGCTCTATCCCAATCTCTGGGAACGTCAGCTTCTACAGCTGTGCAGGGGGCGGCTG 1361  
 QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454  
 Db 1362 TTCCAGCAAGCTGCCGACAGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1421  
 RESULT 14  
 AAT89181  
 ID AAT89181 standard; cDNA; 1768 BP.  
 XX  
 AC AAT89181;  
 XX  
 DT 27-APR-1998 (first entry)  
 XX  
 DE Human chitotriosidase variant cDNA.  
 XX  
 KW Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;  
 XX  
 KW rheumatoid arthritis; atherosclerosis; human; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 124..1524  
 FT CDS /\*tag= a  
 XX  
 XX WO9736917-A1.  
 XX  
 PD 09-OCT-1997.  
 XX  
 PF 21-MAR-1997; 97WO-US05072.  
 XX  
 PR 29-MAR-1996; 96US-0014295.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (SMIK) SMITHKLINE BEECHAM CORP.  
 XX  
 PI Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;  
 XX  
 DR WPI; 1997-503041/46.  
 DR P-PSDB; AAW31498.  
 XX  
 PT New isolated human chitotriosidase gene - used to develop products  
 PT for the diagnosis and treatment of tissue remodeling disorders, e.g.  
 PT rheumatoid arthritis  
 XX  
 PS Claim 1; Page 21-22; 34pp; English.  
 XX  
 CC This human cDNA sequence encodes chitotriosidase (see AAW31498). It  
 CC is a preferred variant of the chitotriosidase cDNA sequence given  
 CC in AAT89180. Also claimed are: (1) a method of diagnosing a tissue  
 CC remodelling disorder related to expression of a mutated

CC chitotriosidase protein in a host comprising carrying out nucleic  
 CC acid amplification; and (2) a method of detecting altered  
 CC expression of a chitotriosidase protein in a host comprising  
 CC contacting a bodily sample with an antibody. Chitotriosidase can  
 CC degrade extracellular matrix substrates with a similar carbohydrate  
 CC structure to chitin. It can be used to develop products which can  
 CC be used in the diagnosis and treatment of tissue remodelling  
 CC disorders such as rheumatoid arthritis or atherosclerosis.  
 CC Chitotriosidase nucleic acids can be used in the recombinant  
 CC production of the enzyme, and as probes or primers in diagnostic  
 CC assays.  
 XX

SQ Sequence 1768 BP; 401 A; 529 C; 474 G; 363 T; 1 other;

#### Alignment Scores:

Pred. No.: 1,14e-110 Length: 1768  
 Score: 1316.00 Matches: 239  
 Percent Similarity: 67.17% Conservative: 70  
 Best Local Similarity: 51.96% Mismatches: 127  
 Query Match: 52.51% Indels: 24  
 DB: 18 Gaps: 3

US-10-004-219B-14 (1-455) x AAT89181 (1-1768)

QY 2 GlnLeuThrCysTyrrPheThrAsnTrpAlaGlnTyrrArgProGlyLeuGlyArgPheMet 21  
 Db 190 AAACCTGGTCTGTACTTCCACAACTGGGCCCCAGTACAGACAGGGGGAGGCTCGCTTCCTG 249  
 QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrrAlaPheAlaGlyArgGln 41  
 Db 250 CCCAAGGACTTGGACCCCGAGCCTTTGACACCACCTCATCTAGCCCTTCGCTGGCATGACC 309  
 QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrrGlnAlaPheAsnGly 61  
 Db 310 AACCCAGCTGAGCACCACCTGAGTGAATGACGAGACTCTTACCAGAGTTCAATGGC 369  
 QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81  
 Db 370 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGAGGCTGGAATTTTCRCG 429  
 QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101  
 Db 430 ACTCAGAAGTTCACAGATATGTTAGCCACGCCCAACACCGTCAGACCTTTGTCAACTCG 489  
 QY 102 ValIleLysPheLeuArgGlnTyrrGluPheAspGlyLeuAspPheAspTrpGluTyrrPro 121  
 Db 490 GCCATCAGTTTCTCGGCAATACAGCTTTGACGGCCTTGACCTTGATGGGAGTACCCA 549  
 QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141  
 Db 550 GGAAGCCAGGGAGGCCCTGCGGTAGACAAGGAGCGCTTCACAACCCCTGGTACAGACTTG 609  
 QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161  
 Db 610 GCCAATGCCCTCCAGCAGGAAGCCACCTCAGGGAAGGAAACGCCCTTCTCTGAGTGCA 669  
 QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrrGluIleProGlnLeuSerGln 181  
 Db 670 GGGTTCCAGCTGGCAGACCTATGTGATGTAGTACGAGGTGGACAAAATGCCCCG 729  
 QY 182 TyrLeuAspTyrrIleHisValMetThrTyrrAspLeuHisGlySerTrpGluGlyTyrrThr 201  
 Db 730 AACCTGGATTTTGTCAACCTTATGGCCTACGACTTCCATGGCTTTGGGAGAAGGTCACG 789  
 QY 202 GlyGluAsnSerProLeuTyrrLysTyrrProThrAspThrGlySerAsnAlaTyrrLeuAsn 221  
 Db 790 GGACATAACAGCCCTCTTACAAGAGGCAAGAGAGTGGTGACGACGCCCTCCAC 849  
 QY 222 ValAspTyrrValMetAsnTyrrTriplysAspAsnGlyAlaProAlaGluLysLeuIleVal 241  
 Db 850 GTGGATGTCTGTGCAACAGTGGCTGCGAGAAGGGGCCCTCGCCAGCAGCTGATCCTT 909  
 QY 242 GlyPheProThrTyrrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261

Db 910 GGCATGCTTACCTACGAGGCTCTCTTACACATGGGCTCTCATCAGACACACAGAGTGGG 969  
 Qy 262 AlaProThrSerGlyValGlyProAlaGlyProTyrAlaValGlySerGlyIleTrpAla 281  
 Db 970 GGGCCAGCCAGGAGTCTGGCTCTCAGGCCCCCTTCCACCAAGGAAGAGGAGTCTGGCC 1029  
 Qy 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301  
 Db 1030 TACTATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGAAATCAGGATCAG 1083  
 Qy 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321  
 Db 1084 AAGGTGCCCTACATCTTCGGGACCAACCAAGTGGTGGCTTTTGATCATGTGGAGAGCTTC 1143  
 Qy 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyValMetValTrpAlaIle 341  
 Db 1144 AAACCAAGGTGAGTATCTGAAGCAGAAAGGAGTGGGGCGGCGCATGCTCTGGGACTG 1203  
 Qy 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361  
 Db 1204 GACTTAGATGACTTTGGCGGCTTCTCTGCAACCAAGGCGCGATACCCCTCATCCAGAGC 1263  
 Qy 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374  
 Db 1264 CTACGGCAGGAAGTCTGCTTCCATCTTGCCTTCCAGGCAACCCAGAGCTTGAAGTTCCA 1323  
 Qy 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394  
 Db 1324 AAACGAGTCAAGCTCTGAACCT----- 1347  
 Qy 395 GlySerSerSerGlyGlySerGlySerGlyPheCysAlaValArgAlaAsn 414  
 Db 1348 -----GAGCATGGCCCGCCAGCTCTGCAAGACACGCTTCTGCCAGGGCAAGCTGAT 1398  
 Qy 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434  
 Db 1399 GGGCTCTATCCCAATCTCTGGGAACGGTCTCAGCTTCTACAGCTGTGACGCGGGGGCTG 1458  
 Qy 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454  
 Db 1459 TTCCAGCAAGCTGCCCGACAGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1518

RESULT 15  
 AAT50834  
 ID AAT50834 standard; cDNA; 1713 BP.  
 XX  
 AC AAT50834;  
 XX  
 DT 24-MAR-1997 (first entry)  
 XX  
 DE Human chitinase cDNA clone chi.39.  
 XX  
 KW Chitinase; chitotriosidase; chitin; infectious disease;  
 KW gene therapy; vaccine; lysosomal lipidoses; Gaucher's disease;  
 KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;  
 KW multiple sclerosis; drug delivery; cosmetics; food; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT CDS 13..1176  
 FT /\*tag= a  
 FT sig\_peptide 13..75  
 FT /\*tag= b  
 FT mat\_peptide 76..1173  
 FT /\*tag= c  
 XX  
 PN WO9640940-A2.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 06-JUN-1996; 96WO-NL00225.

XX 07-JUN-1995; 95US-0468839.  
 PR (UNAM ) UNIV AMSTERDAM.  
 XX  
 PA Aerts JMG;  
 PI  
 XX WPI; 1997-118698/11.  
 DR P-PSDB; AAW08585.  
 XX  
 PT New human chitinase - used to treat or prevent infection by  
 PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,  
 PT foods, implants etc.  
 XX  
 PS Claim 2; Page 42-43; 58pp; English.  
 XX  
 CC A cDNA clone (AAT50834), designated chi.39, codes for a 39 kDa human  
 CC chitinase (AAW08584) that is stable to many proteases, active at pH 3-  
 CC 8 and up to 50 deg, and stable in the circulation. Clones chi.39  
 CC and chi.50 (see also AAT50833) were isolated from a human macrophage  
 CC cDNA library using as probe a partial clone obtd. using primers  
 CC (see also AAT50835-36) based on a chitotriosidase purified from a  
 CC type 1 Gaucher disease patient. The 2 cDNA clones are the result  
 CC of alternative splicing of RNA. Chitinase nucleic acid can be used  
 CC for large-scale prodn. of recombinant human chitinases, or can be  
 CC incorporated into a gene therapy vector to treat or prevent  
 CC infection by chitin-contg. pathogens.  
 XX  
 SQ Sequence 1713 BP; 380 A; 503 C; 465 G; 365 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,82e-110 Length: 1713  
 Score: 1311.50 Matches: 244  
 Percent Similarity: 68.03% Conservative: 76  
 Best Local Similarity: 51.91% Mismatches: 131  
 Query Match: 52.33% Indels: 19  
 DB: 18 Gaps: 4  
 US-10-004-219b-14 (1-455) x AAT50834 (1-1713)  
 Qy 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21  
 Db 79 AAACCTGGTCTGCTACTTCCACCACTGGGCCCATGACAGAGGGAGGCTCGCTTCCTG 138  
 Qy 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41  
 Db 139 CCCAAGGACTTGGACCCAGCCAGCTTTGACCCACCTCATCTACGCTTCGCTGGCATGACC 198  
 Qy 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61  
 Db 199 AACCAACAGCTGAGCACCACCTGAGTGAATGACGAGACTCTCTACAGGAGTTCAATGSC 258  
 Qy 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyIleGlyTrpAsnPheGly 81  
 Db 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTTGTAGCCATCGAGGCTGGAAATTCGGC 318  
 Qy 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101  
 Db 319 ACTCAGAAGTTTACAGATATGTTAGCCAGCGGCAACACCGTCAGACCTTTGTCAACTCG 378  
 Qy 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121  
 Db 379 GCCATCAGTTTCTCGGCAATATACAGCTTTGACGGCTTGTACCTTGGAGTACCCA 438  
 Qy 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141  
 Db 439 GGAAGCCAGGGAGGCCCTTCGGGTAGACAAGGAGCGCTTCACACACCTTGTGTAGAGACTTG 498  
 Qy 142 ArgGluAlaPheGluGlnAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161  
 Db 499 GCCAATGCCCTCCAGGAGAACCCAGCTCAGGGAAGAACGCTTCTTCTGATGCA 558  
 Qy 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181

[illegible]

Search completed: June 29, 2003, 21:27:24  
Job time : 271.712 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:12:38 ; Search time 64.9519 Seconds

(without alignments)

2148.326 Million cell updates/sec

Title: US-10-004-219B-14

Perfect score: 2506

Sequence: 1 YQLTCYFTNWAQYRPLGRF.....QQNCOAGLVFTSCDCNWA 455

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO spool/US10004219/runat\_24062003\_160229\_8184/app\_query.fasta\_1.1877  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CGN 1 1 108 @runat\_24062003\_160229\_8184 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1323	52.8	1636	US-09-039-198A-1	Sequence 1, Appli
2	1323	52.8	1636	US-08-877-599-1	Sequence 1, Appli
3	1323	52.8	1636	US-09-267-574-1	Sequence 1, Appli
4	1323	52.8	1643	US-08-486-839-3	Sequence 3, Appli
5	1323	52.8	1643	US-09-151-011-3	Sequence 3, Appli
6	1323	52.8	1643	US-09-343-623-3	Sequence 3, Appli
7	1317	52.6	1656	US-09-039-198A-3	Sequence 3, Appli
8	1317	52.6	1656	US-08-877-599-3	Sequence 3, Appli
9	1317	52.6	1656	US-09-267-574-3	Sequence 3, Appli
10	1311.5	52.3	1713	US-08-486-839-5	Sequence 5, Appli
11	1311.5	52.3	1713	US-09-151-011-5	Sequence 5, Appli
12	1311.5	52.3	1713	US-09-343-623-5	Sequence 5, Appli

13	970.5	38.7	1433	1	US-08-694-915-1	Sequence 1, Appli
14	970.5	38.7	1526	1	US-08-694-915-3	Sequence 3, Appli
15	962.5	38.4	1681	5	PCT-US94-07754-4	Sequence 4, Appli
16	957.5	38.2	1681	2	US-08-581-527-4	Sequence 4, Appli
17	863	34.4	2452	2	US-08-524-051-1	Sequence 1, Appli
18	798.5	31.9	1677	4	US-09-545-814-13	Sequence 13, Appli
19	798.5	31.9	1677	4	US-09-545-814-15	Sequence 15, Appli
20	798.5	31.9	1749	4	US-09-545-814-4	Sequence 4, Appli
21	798.5	31.9	1749	4	US-09-545-814-6	Sequence 6, Appli
22	798.5	31.9	2610	4	US-09-545-814-1	Sequence 1, Appli
23	798.5	31.9	2610	4	US-09-545-814-3	Sequence 3, Appli
24	794.5	31.7	1919	4	US-09-545-814-31	Sequence 31, Appli
25	794.5	31.7	1919	4	US-09-545-814-33	Sequence 33, Appli
26	711.5	28.4	1470	4	US-09-292-225-40	Sequence 40, Appli
27	711.5	28.4	1470	4	US-09-292-225-42	Sequence 42, Appli
28	711.5	28.4	1527	4	US-09-292-225-37	Sequence 37, Appli
29	711.5	28.4	1527	4	US-09-292-225-39	Sequence 39, Appli
30	711.5	28.4	1621	4	US-09-292-225-34	Sequence 34, Appli
31	711.5	28.4	1621	4	US-09-292-225-36	Sequence 36, Appli
32	708.5	28.3	1608	4	US-09-292-225-20	Sequence 20, Appli
33	708.5	28.3	1608	4	US-09-292-225-22	Sequence 22, Appli
34	708.5	28.3	1665	4	US-09-292-225-17	Sequence 17, Appli
35	708.5	28.3	1665	4	US-09-292-225-19	Sequence 19, Appli
36	708.5	28.3	1752	4	US-09-292-225-14	Sequence 14, Appli
37	708.5	28.3	1752	4	US-09-292-225-16	Sequence 16, Appli
38	672	26.8	1478	4	US-09-545-814-28	Sequence 28, Appli
39	672	26.8	1478	4	US-09-545-814-30	Sequence 30, Appli
40	472.5	18.9	1167	1	US-07-939-501A-6	Sequence 6, Appli
41	472.5	18.9	1167	4	US-08-448-398-10	Sequence 10, Appli
42	472.5	18.9	1320	1	US-07-939-501A-14	Sequence 14, Appli
43	472.5	18.9	1364	1	US-07-939-501A-13	Sequence 13, Appli
44	472.5	18.9	1405	1	US-07-939-501A-9	Sequence 9, Appli
45	463.5	18.5	1554	1	US-08-045-269C-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-039-198A-1  
; Sequence 1, Application US/09039198A  
; Patent No. 6200951  
; GENERAL INFORMATION:  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Tjoelker, Larry W.  
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,198A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27866/34391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1636 base pairs  
; TYPE: nucleic acid

STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 2..1399  
 FEATURE:  
 NAME/KEY: mat\_peptide  
 LOCATION: 65..1399  
 US-09-039-198A-1

Alignment Scores:  
 Pred. No.: 6,166-123 Length: 1636  
 Score: 1323.00 Matches: 240  
 Percent Similarity: 67.39% Conservative: 70  
 Best Local Similarity: 52.17% Mismatches: 126  
 Query Match: 52.79% Indels: 24  
 DB: 4 Gaps: 3

US-10-004-219B-14 (1-455) x US-09-039-198A-1 (1-1636)

```

QY      2 GlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB      68 AAACCTGGTCTGCTACTTCCAACTGGGCCCTGATACAGACAGGGGAGGCTGCTTCCTG 127

QY      22 ProAspAsnLeuAspProCysLeuCysThrHisLeuLeuTyrAlaPheAlaGlyArgGln 41
DB      128 CCCAAGGACTTGACCCCGCCAGCTTTCGCCACCTCATCTACGCCCTTCGCTGGCATGACC 187

QY      42 AsnAsnGlnLeuThrThrLeuGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB      188 AACCCACGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTTACACGAGGATTCATATGGC 247

QY      62 LeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyClyTrpAsnPheGly 81
DB      248 CTGAAGAAGATGAATCCCAAGCTGAGAGCCCTGTAGCCATCGAGGCTGGAATTTCCGCG 307

QY      82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
DB      308 ACTCAGAAGTTACACATATGTTAGCCAGCCGACCAACCGTCAGACCTTTGTCAACTCG 367

QY      102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB      368 GCCATCAGGTTCTCGCGAAATACAGCTTTGACGGCTTGACCTTGAGTGGAGTACCCA 427

QY      122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB      428 GGAAGCCAGGGAGCCCTGCCGTAGCAAGGAGCGCTTCACACCTCGGTACAGGACTTG 487

QY      142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
DB      488 GCCAATGCCCTTCAGCAGGAGGAGCCAGACTCAGGAGAGGAGCCCTTCTTCGAGTGA 547

QY      162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
DB      548 GCGGTTCCAGCTGGGAGCAGACTATGTTGGATGCTGGATACGAGTGACAAATCGGCCAG 607

QY      182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB      608 AACCTGGATTTGTCAACCTTATGGCTCAGACTTCCATGGCTCTTGGGAGAGGTCACG 667

QY      202 GlyGluAsnSerProLeuTyrIleLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB      668 GGCATACAGCCCTCTACAGAGGCGCAGAGAGAGTGGTGCACGACCCAGCCTCAAC 727

QY      222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
DB      728 GTGGATGCTGCTGTGCAACAGTGGCTGTCAGAGGAGGAGCCCTGCCAGCAAGCTGATCCTT 787

QY      242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
DB      788 GGCATGCCCTACCTACGAGCGCTCTTTCACACTGGCTCTCTCATCAGACACAGAGTGGG 847
  
```

```

QY      262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
DB      848 GCCCAGCCACAGGCTTGGCACTCCAGGCCCTTCCACCAAGGAAGGAGGATGCTGGCC 907

QY      282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
DB      908 TACTATGAAGTCTGCTCTGG-----AAGGGGGGCCCAACACAGAGAATCCAGGATCAG 961

QY      302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
DB      962 AAGGTGCCCTCATCTTCGGGACACACAGTGGGTTGATGATGATGATGAGAGCTTC 1021

QY      322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyValaMetValTrpAlaIle 341
DB      1022 AAAACCAAGGTCACTATCTGAAGCAGAGGACTGGGGGGGCCCATGTCTGGGCACTG 1081

QY      342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
DB      1082 GACTTAGATGACTTTGCGCGCTTCTCTGCAACACAGGCGCATACCCCTCATCCAGAGC 1141

QY      362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
DB      1142 CTACGGCAGGAACCTGAGTCTTCATCTTCAGGCCACCCAGAGCTTGAAGTTCCA 1201

QY      375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB      1202 AAACCAAGGTCAAGCTCTGAACCT----- 1225

QY      395 GlySerSerSerGlyClySerSerGlyGlySerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
DB      1226 -----GAGCATGGGCCCGCCAGCCCTGACAAAGACACGTTCTGCGCAGGGCAAGCTGAT 1276

QY      415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValaAsnGlyValThr 434
DB      1277 GGGCTCTATCCCAATCTCGGGAAACGGTCCAGCTCTACAGCTGTGCGAGCGGGGGCTG 1336

QY      435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
DB      1337 TTCCAGCAAGAGTGGCCCGACAGGCTGTGTGTTTCAGCAACTCTCTGCAATGCTGCACCTGG 1396

RESULT 2
US-08-877-599-1
; Sequence 1, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
  
```



## Alignment Scores:

Pred. No.: 6,156-123 Length: 1636  
Score: 1323.00 Matches: 240  
Percent Similarity: 67.39% Conservative: 70  
Best Local Similarity: 52.17% Mismatches: 126  
Query Match: 52.79% Indels: 24  
DB: 4 Gaps: 3

US-10-004-219B-14 (1-455) x US-09-267-574-1 (1-1636)

```
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 68 AAATGGTCTGCTTACTTCAACCACTGGGCCAGTACAGACAGGGGAGGCTGCTTCCTG 127
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleThrAlaPheAlaGlyArgGln 41
DB 128 CCAGAGACTTGGACCCCGACCTTTGCACCCACCTCATCTACGCCCTTGCCTGGCATGACC 187
QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 188 ACCACCACTGAGCACCAGTCTGAGTGGATGACGACTCTTACACGAGGTTCAATGGC 247
QY 62 LeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
DB 248 CTGAAGAAGATGAATCCCAAGCTGAAGACCTCTTAGCCATCGAGGCTGGAATTTCCGC 307
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
DB 308 ACTCAGAAGTTTACAGATATGGTAGCGCCGCAACACCGTCAGACCTTTGTCAACTCG 367
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 368 GCCATCAGGTTCTGCGCAATACAGCTTTGACGGCTTGACCTTCTGAGTACCCCA 427
QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 428 GGAAGCAGGGAGCGCTGCGGAGCAAGGAGCGCTTCAACACCTTGTGTACAGGACTTG 487
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
DB 488 GCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGGAACGCTTCTTCTGAGTGCA 547
QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
DB 548 CGGTTCCAGTGGCGCAGACCTATGCGATGCTGGATACGAGTGACAAATCGCCAG 607
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 608 AACCTGGATTTGTCAACTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTCAAG 667
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 668 GGACATAACAGCCCTCTCACAAGAGCGCAAGAGAGTGGTGCAGCAGCCAGCCCTCAAC 727
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
DB 728 GTGATGCTCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTT 787
QY 242 GlyPheProThrTyrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyIleGly 261
DB 788 GGCATGCTTACCTTACGAGGCTCTTCACTGGCTCTCATCAGACACCAAGAGTGGGG 847
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
DB 848 GCCCCAGCACAGGGTCTGCGACTCCAGGCCCTTCCAAAGGAAGAGGATGCTGGCC 907
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyValaThrGlnGlyTrpAspAlaProGln 301
DB 908 TACTATGAAGTCTGCTCTCG-----AAGGGGGCCCAACACAGAGATCCAGATCAG 961
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
DB 962 AAGGTGCCCTTACATCTTCCGGGCAACACCAAGTGGGCTTTGATGATGTGGAGAGCTTC 1021
```

```
QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
DB 1022 AAAACCAAGCTCAGCTATCTGAAGCAGAGGACTGGGGGGGCCCATGCTTGGGCACTG 1081
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
DB 1082 GACTTAGATGACTTTTGGCGGCTTCTCTGCAACCAAGGCGCATACCCCTCATCCAGACG 1141
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
DB 1142 CTACCGCAGGAAGTCACTGCTTCCATCTTCCAGGCACCCAGAGCTTGAAGTTCCA 1201
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB 1202 AAACCAAGTCAAGCTCTGAACT----- 1225
QY 395 GlySerSerSerSerGlyLysSerGlySerGlySerGlyPheCysAlaValAlaArgAla 414
DB 1226 -----GAGCATGGCCCCAGCCCTGGACAAGACACGTTCTGCCAGGGCAAGCTGAT 1276
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
DB 1277 GGGCTCTATCCCAATCCTCGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTG 1336
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
DB 1337 TTCCAGCAAGTGGCCGACAGGCTTGTGTTCAGCAACTCTCTGCAATGCTGCACCTGG 1396
```

## RESULT 4

US-08-486-839-3  
Sequence 3, Application US/08486839  
Patent No. 5928928  
GENERAL INFORMATION:

APPLICANT:  
TITLE OF INVENTION: A human chitinase, its recombinant  
production, its use for decomposing chitin, its use  
in therapy or prophylaxis against infection diseases.  
TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann & Baron  
CITY: Jericho  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11758

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/486,839  
FILING DATE: 07 - June - 1995

## ATTORNEY/AGENT INFORMATION:

NAME: Baron, Ronald J.  
REGISTRATION NUMBER: 29,281  
REFERENCE/DOCKET NUMBER: 294-26  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1643 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

US-08-486-839-3

Alignment Scores:



Pred. No.: 6.2e-123 Length: 1643  
Score: 1323.00 Matches: 240  
Percent Similarity: 67.39% Conservative: 70  
Best Local Similarity: 52.17% Mismatches: 126  
Query Match: 52.79% Indels: 24  
DB: 2 Gaps: 3

US-10-004-219B-14 (1-455) x US-08-486-839-3 (1-1643)

```
QY 2 GlnLeuThrCysThrPheThrAsnTrpAlaGlnTrpArgProGlyLeuGlyArgPheMet 21
DB 79 AAACCTGGCTGCTACTTACCAACTGGGCCAGTACAGACAGGGAGGCTCGCTTCCTG 138
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
DB 139 CCAAGGACTGGACCCAGCCTTTGACCCACCTCATCTACGCTTCTGCTGGCATGCC 198
QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 199 AACCAACAGCTGGACCACTGAGTGGATGACGAGACTCTTACCAAGAGTTCAATGCC 258
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyClyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTAGCCATCGAGGCTGGAAATTCGSC 318
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
DB 319 ACTCAGAAGTTCACAGATATGTTAGCCAGCGGCCCAACACCGTCAGACCTTTGTCAACTCG 378
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTCTCGCAATACAGCTTGTAGCGCTTTGACCTTGTGAGTACCCA 438
QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 439 GGAACCCAGGGAGCCCTCCCTAGACAGAGAGCGCTTCAACACCTGTGTACAGACTTG 498
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
DB 499 GCCAATGCTTCCAGCAGGAGGCCAGCCTCAGGAGGAGGAGCGCTTCTTCTGAGTGCA 558
QY 162 AlaValAlaAlaGlyLysSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
DB 559 GCGGTTCCAGCTGGGACAGCACTATGTGATGCTGGATACGAGTGGACAAATATGCCAG 618
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 619 AACCTGGATTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTCCAG 678
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 679 GGACATAACAGCCCTCTACAGAGGCAAGAGAGTGGTGACAGCAGCCTCAAC 738
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuVal 241
DB 739 GTGGATGCTGCTGTGCACAGTGGCTGCAGAGGGAGCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
DB 799 GGCATGCTTACCTAGGAGCGCTTCTACACTGGGCTCTCATCAGACACAGAGTGGGG 858
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
DB 859 GCGCCAGCACAGGCTGCGCACTCCAGCCCTTCCACCAAGAGAGGAGGATGCGGCC 918
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGln 301
DB 919 TACTATGAAGTCTGCTCCTGG-----AAGGGGGGCCACCAACAGAGAATCCAGATCAG 972
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
DB 973 AAGGTGCCCTATCTTCCGGGACCAACAGTGGGTGGGCTTTGATGATGTGGAGAGCTTC 1032
```

```
QY 322 AspleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
DB 1033 AAAACCAAGGTGAGTATCTGAAGCAGAGGAGTGGGGGGCCATGCTCTGGGCACTG 1092
QY 342 AspleuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
DB 1093 GACTTAGATGACTTTGGCGGCTTCTCTGCAACCAAGGCGCCATACCCCTCATCCAGAGC 1152
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
DB 1153 CTAGGAGGAGTACTGCTTCCATCTTGGCTTCCAGCACCCAGAGCTTGAAGTTCCA 1212
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB 1213 AAACGAGTTCAGCCCTCTGAACCT----- 1236
QY 395 GlySerSerSerSerGlySerSerGlySerGlySerGlySerGlySerGlySerGlySer 414
DB 1237 -----GAGCATGGGCCCCAGCGCTTGTCCAGGCGCAAGCTGAT 1287
QY 415 GlyLeuTyrProValAlaAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
DB 1288 GGGCTCTATCCCAATCTCTCGGAGAGCTTCCAGCTTCTACAGCTGTGCAGCGGGGGCTG 1347
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysAsnTrp 454
DB 1348 TTCAGCAAGCTGCCGACAGGCTGTGTTCAGCAACTCTCTGCAATGCTGCACCTGG 1407
```

## RESULT 5

US-09-151-011-3  
; Sequence 3, Application US/09151011  
; Patent No. 6057142

## GENERAL INFORMATION:

APPLICANT:  
TITLE OF INVENTION: A Human Chitinase, Its Recombinant  
TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Baron, LLP  
CITY: Syosset  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11791

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/151,011  
FILING DATE: 10 - September - 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Morris, Robert C.  
REGISTRATION NUMBER: 42,910  
REFERENCE/DOCKET NUMBER: 294-32 DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO

US-09-151-011-3

Alignment Scores: 6.2e-123 Length: 1643  
Pred. No.: 1643

Score: 1323.00 Matches: 240  
Percent Similarity: 67.39% Conservative: 70  
Best Local Similarity: 52.17% Mismatches: 126  
Query Match: 52.79% Indels: 24  
DB: Gaps: 3

US-10-004-219B-14 (1-455) x US-09-151-011-3 (1-1643)

```
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
Db 79 AAACGTGGTCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 138
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
Db 139 CCAGAGGACTTGGACCCAGCCCTTTCACCCACCTCATCTAGCCCTTCGCTGGCATGACC 198
QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
Db 199 AACCCACGCTGAGCACCCTAGTGGATGAGATGACGAGACTCTTACAGGAGTTCAATGGC 258
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81
Db 259 CTGAAGAAGTGNATCCCAAGCTGAGACCCCTTATGACCATCGGAGCTGGAAATTCGGC 318
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
Db 319 ACTCAGAAGTTCACAGATATGTTAGTCAGCGGCCCAACACCGTCAGACCTTTGTCAACTCG 378
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
Db 379 GCATCAGGTTCTGCGCAAAATACACTTTCAGCGCTTTCAGCTTTCAGTGGAGTACCCA 438
QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
Db 439 GGAAGCCAGGGAGCCCTCCGTAGACAGGAGCGCTTCACACCTGTTGACAGACTTG 498
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
Db 499 GCCAATGCTCTCCAGCAGGAAGCCAGACCTCAGGGAAGAAAGCGCTTCTTCTGAGTGA 558
QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
Db 559 GCGGTTCAGCTGGCAGACCTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 618
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
Db 619 AACCTGGATTTGTCACCTTATGCTGCTAGCTTCCATGGCTTTCGAGAGGCTCAGC 678
QY 202 GlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
Db 679 GGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGCAGCCAGCTCAAC 738
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
Db 739 GTGATGCTGCTGTGCAACAGTGGCTGCAAGAGGAGCCCTGCGCAGAGCTGATCCTT 798
QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
Db 799 GGCATGCTTACCTACGAGCGCTCTTCACACTGGCTCTTCATCAGACACAGAGTGGG 858
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
Db 859 GCCCCAGCCACAGGCTGTCGCACTCCAGGCCCCCTTCACCAAGGAGGAGGAGTGGGCC 918
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
Db 919 TACTATGAGTCTGCTGCTG-----AGGGGGGCCCAACACAGAGATCCAGGATCAG 972
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
Db 973 AAGTGCCCTACATCTCCGGGACACACAGTGGTGGCTTTGATGATGTGGAGCTTC 1032
QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
```

```
Db 1033 AAAACAAGGTCAGCTATCTGAAGCAGAGGAGTGGGGGGCCCATGCTCTGGGCACTG 1092
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
Db 1093 GACTTAGATGACTTTCGCGGCTTCTCCTGCAACAGGCGCCGATACCCCTCATCCAGAG 1152
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
Db 1153 CTACGGCAGGAACCTAGTCTTCCATACTTGCCTTCAGGCACCCAGAGCTTGAAGTTCCA 1212
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
Db 1213 AAACAGGTTCAGCCCTCTGAACCT----- 1236
QY 395 GlySerSerSerSerGlyLysSerSerGlyGlySerGlyPheCysAlaValAlaArgAlaAsn 414
Db 1237 -----GAGCATGGCCCCAGCCCTGCAGACAGACGTTCTGCAGGGCAAGCTGAT 1287
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
Db 1288 GGGCTTCTATCCCAATCCTCGGGAAGCGTCCAGCTTCTACAGCTGTGCAGCGGGGGCTG 1347
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
Db 1348 TTCCAGCAAGCTGCCGACAGGCTGGTGTTCAGCAACTCTGCAAAATGCTGCACCTGG 1407

RESULT 6
US-09-343-623-3
; Sequence 3, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-09-343-623-3
```

NAME/KEY: mat peptide

LOCATION: 90..1424  
US-09-039-198A-3

## Alignment Scores:

Pred. No.: 2, 5e-122 Length: 1656  
Score: 1317.00 Matches: 239  
Percent Similarity: 67.17% Conservative: 70  
Best Local Similarity: 51.96% Mismatches: 127  
Query Match: 52.55% Indels: 24  
DB: 3 Gaps: 3

US-10-004-219B-14 (1-455) x US-09-039-198A-3 (1-1656)

```
QY 2 GlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 93 AAACGTGCTGCTACTTCCACCACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 152
QY 22 ProAspAsnLeuAspProCysLeuCysThrHisLeuLeuLeuLeuLeuLeuLeuLeu 41
DB 153 CCCAAGGACTTGGACCCAGCCCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATGACC 212
QY 42 AsnAsnGluLeuThrThrLeuTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 213 AACCCAGCTGAGCAGCACTGAGTGGATGAGAGACTCTTACCAGAGATTCAATGGC 272
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaLeuLeuLeuLeuLeuLeu 81
DB 273 CTGAAGAAGTGAATCCCAAGCTGAGAGCCCTGTAGCCATCGGAGCTTGAATTCAGC 332
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheLeuThrSer 101
DB 333 ACTCAAGATTCACAGATATGTTAGTGGTGGAGCCCAACAGCTGAGACCTTTGTCAACTCG 392
QY 102 ValLeuLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 393 GCCATCAGGTTTCTGGCGAAATACACTTTCAGCGCTTACCTGAGTGGAGTACCA 452
QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 453 GGAAGCCAGGGAGCCCTCGCTGAGCAAGAGGAGCGCTTCACAACTGTTGACAGGACTG 512
QY 142 ArgGluAlaPheGluGlnGlnAlaLysGlnLeuAsnLysProArgLeuMetValThrAla 161
DB 513 GCCAATGCTCTCCAGCAGGAAGCCCAAGCTCAGGGAAGGAAAGCGCTTCTCTGAGTGCA 572
QY 162 AlaValAlaAlaGlyLeuSerHisLeuGlnSerGlyTyrGluLeuProGlnLeuSerGln 181
DB 573 GCGGTTCCAGCTGGGAGCACTATGATGGATGCTGGATAGAGTGGACAAATTCGCCAG 632
QY 182 TyrLeuAspTyrLeuHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 633 AACCTGGATTGTCACCTTATGGCTTACGACTTCCATGGCTTCTGGGAGAGGTCAGC 692
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 693 GGACATAACAGCCCCCTCTACAAGAGGCAAGAGAGTGGTGCAGCAGCAGCCTCAAC 752
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyValProAlaGluLysLeuLeuVal 241
DB 753 GTGGATGCTGTGCAACAGTGGCTGCGAGAGGGAGCCCTGCCAGCAAGCTGATCCTT 812
QY 242 GlyPheProThrTyrGlyHisAsnPheLeuLeuSerAsnProSerAsnThrGlyLeuGly 261
DB 813 GGCATGCTTACCTACGAGCGCTCTTACACTGGCTCTCATCAGACACAGAGTGGGG 872
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyLeuThrAla 281
DB 873 GCCCCAGCACAGGGTCTGGCACTTCCAGGCCCTTCCACCAAGGAGGAGGATGTGGCC 932
QY 282 TyrTyrGluLeuLysCysThrPheLeuLysAsnGlyValThrGlnGlyTrpAspAlaProGln 301
DB 933 TACTATGAAGTCTGCTCTGG-----AAGGGGGGCCCAACAGAGAAATCCAGGATCAG 986
```

```
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnLeuLysSerPhe 321
DB 987 AAGGTGGCCCTACATCTTCGGGACCAACAGTGGTGGGCTTTGATGATGTGGAGAGCTTC 1046
QY 322 AspLeuLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaLeu 341
DB 1047 AAACCAAGAGTTCAGTATCTGAAGCAGAGGAGCTGGGGGGGGCCATGCTCTGGGCACTG 1106
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuLeuSerThr 361
DB 1107 GACTTAGATGACTTTCGGGCTTCTCTGCAACAGGCGCCGATACCCCTCATCCAGAGC 1166
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
DB 1167 CTACGGCAGGAAGTCTTCCATCTTCATCTTCAGGCACCCAGAGCTTCAAGTTCCA 1226
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB 1227 AAACGAGTTCAGCCCTCTGAACCT----- 1250
QY 395 GlySerSerSerSerGlyGlySerGlyGlySerGlyGlyPheCysAlaValArgAlaAsn 414
DB 1251 -----GAGCATGGCCCCAGCCCTGGACAGACAGCTTTCGCCAGGGCAAGCTGAT 1301
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
DB 1302 GGGCTCTATCCCAATCTCTCGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGGGCTG 1361
QY 435 TyrGlnGlnAsnGlyGlnAlaGlyLeuValPheAspThrSerCysAspCysValAsnTrp 454
DB 1362 TTCCAGCAAGAGTCCCGCAGAGCGCTGTGTGTTTCAGCAACTCTCTGCAAAATGTCACCTGG 1421

RESULT 8
US-08-877-599-3
; Sequence 3, Application US/08877599
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

Alignment Scores:		
Pred. NO.:	2.5e-122	Length:
Score:	1317.00	Matches:
Percent Similarity:	67.17%	Conservative:
Best Local Similarity:	51.96%	Mismatches:
Query Match:	52.55%	Indels:
DB:	4	Gaps:
		3
		1556



Qy	362	LeuYsGlyAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnPro-IleG1	381
Db	1153	CTACGGCAGGAACCTGAATGGGTAAAGCCCTTAACCTGCTCTGCATCATGTGAGGCCAGGTGTT	1212
Qy	381	uProIleThr-AlaAla-----ProSerGlySerGlyAsn-	392
Db	1213	GCCTGTGGCACTGTGCTTCAGCTGTAGTCTTCCATACATTGCTTCAGGCACCCAGAGC	1272
Qy	393	-----GlySerGlySerSerSerSerGlyGlySerSerGlyGlySerG	407
Db	1273	TTGAAGTTCCAAAACACAGGTTCAGCCCTCTGAACCTGAGCATGCGCCCAAGCCCTGGACAAG	1332
Qy	407	ly---PheCysAlaValAlaGlnAlaAsnGlyLeuTyPrProValAlaAlaAsnAsnArgAsnAlap	426
Db	1333	ACAGTTCCTGCCAGGCAAGCTGATGGGCTCTATCCCAATCTTCGGGAACGGTCCAGCT	1392
Qy	426	hetrPhisCysValaAsnGlyValThrTyGlnGlnAsnCysGlnAlaGlyLeuValPheA	446
Db	1393	TCTACAGCTGTGCAGCGGGGGGCTGTTCCAGCAAAAGCTGCCCGACAGGCTTGGTGTTC	1452
Qy	446	spThrSerCysAspCysCysAsnTyr	454
Db	1453	GCAACTCTGCAAATGCTGCACCTGG	1478

RESULT 11

US-09-151-011-5

Sequence 5, Application US/09151011

Patent No. 6057142

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: A Human Chitinase, Its Recombinant

TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use i

TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Barton, LLP

STREET: 6900 Jericho Turnpike

CITY: Syosset

STATE: New York

COUNTRY: United States of America

ZIP: 11791

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

SOFTWARE:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/151,011

FILING DATE: 10 - September - 1998

ATTORNEY/AGENT INFORMATION:

NAME: Morris, Robert C.

REGISTRATION NUMBER: 42,910

REFERENCE/DOCKET NUMBER: 294-32 DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1713 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

US-09-151-011-5

Alignment Scores:

Pred. No.: 9,37e-122 Length: 1713

Score: 1311.50 Matches: 244

Percent Similarity: 68.09% Conservative: 76

Best Local Similarity: 51.91% Mismatches: 131

Length:	1713
Matches:	244
Conservative:	76
Mismatches:	131
Indels:	19

DB: 3 Gaps: 4  
US-10-004-219b-14 (1-455) x US-09-151-011-5 (1-1713)  
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21  
DB 79 AAACCTGGTCTGCTACTTACCAACTGGGCGCCAGTACAGACAGGGGAGGCTGCTTCTTG 138  
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41  
DB 139 CCCAAGCATGTGGACCCAGCCCTTTGCACCCACCTCATCTACGCCCTTCCTGCGATGACC 198  
QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61  
DB 199 AACCCACCATGTGACCACTGAGTGGATGACGAGACTCTCTACAGGAGTTCAATGGC 258  
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyTyrAsnPheGly 81  
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGAGGCTGGAATTTCCGC 318  
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101  
DB 319 ACTCAGAAGTTACAGATATGTAGCCACGCGCCACACCGTCAGACCTTTGTCAACTCG 378  
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121  
DB 379 GCCATCAGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGCGGAGTACCCA 438  
QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141  
DB 439 GGAAGCCAGGGGAGCCCTCGCTAGCAAGAGGAGCGCTTCACACCTGGTACAGGACTTG 498  
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161  
DB 499 GCCAATGCTTCCAGCAAGCCAGACTCAGGAGAGAGCGCTTCTTCGAGTGCA 558  
QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181  
DB 559 GCGGTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATACGAGTGGACAAATCGCCAG 618  
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201  
DB 619 AACCTGGATTTGTCAACCTTATGGCTCAGACTCCATGGCTCTTGGGAGAGGTCACG 678  
QY 202 GlyGluAsnSerProLeuTyrIleValTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221  
DB 679 GGACATACAGCCCTCTACAGAGCGCAAGAGAGTGGTCCAGCAGCAGCCTCAAC 738  
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241  
DB 739 GTGGATGCTGCTGCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798  
QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261  
DB 799 GGCATGCTTACCTACGAGCGCTCTTCACTGGCCCTCTCATCAGACACAGAGTGGGG 858  
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAla 281  
DB 859 GCCCCAGCCAGAGGCTGGCACTCAGGCCCTTCCACAGAGAGAGGAGGATGCTGGCC 918  
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301  
DB 919 TACTATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGAATCCAGGATCAG 972  
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321  
DB 973 AAGGTGCCCTTACATCTTCCGGGCAACCACTGAGTGGGCTTGTATGATGTGGAGACTTC 1032  
QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341  
DB 1033 AAACACAGGTGAGCTATCTGAAGCAGAGAGGAGGACTGGGCGGCGCATGTCTGGGCACTG 1092  
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361

DB 1093 GACTTAGTAGTACTTTGGCCGGCTTCTCTGCAACGAGGCGCATACCCCTCATCCAGACG 1152  
QY 362 LeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnPro-IleGln 381  
DB 1153 CTACGGCAGGAACCTGAATGGTAAAGCCCTTAAGTCCCTGTTCATGTGAGGCCAGGTGT 1212  
QY 381 uproIleThr-AlaAla-----ProSerGlySerGlyAsn- 392  
DB 1213 GCCTGTGGCACTGTCTTCAGCTGTAGTCTTCCATACCTTGCCTTCAGGCCACCCAGAGC 1272  
QY 393 -----GlySerGlySerSerSerGlyGlySerSerGlyGlySerGlyGlySerG 407  
DB 1273 TTGAAGTTCCAAACACGAGTCTGAACTCTGAACTGTGAGCATGGCCCAAGCCCTGGACAAG 1332  
QY 407 Ly---PheCysAlaValAlaGlnAlaAsnGlyLeuTyrProValAlaAlaAsnAsnArgAsnAlap 426  
DB 1333 ACAGTCTTCCAGGCGGCAAGCTGATGGCTATATCCCAATCCTCGGAGACGCTCCAGCT 1392  
QY 426 heTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPheA 446  
DB 1393 TCTACAGCTGTGACGCGGGGGCTGTTCCAGCAAGCTGCCCGACAGCGCTGTGTTC 1452  
QY 446 spThrSerCysAspCysCysAsnTrp 454  
DB 1453 GCAACTCTCTGCAATGCTGCACCTGG 1478  
RESULT 12  
US-09-343-623-5  
; Sequence 5, Application US/09343623  
; Patent No. 6303118  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A human chitinase, its recombinant  
; TITLE OF INVENTION: production, its use for decomposing chitin, its use  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11758  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; SOFTWARE:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/343,623  
; FILING DATE:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,839  
; FILING DATE: 07-June-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baron, Ronald J.  
; REGISTRATION NUMBER: 29,281  
; REFERENCE/DOCKET NUMBER: 294-26  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1713 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
US-09-343-623-5



## Alignment Scores:

Pred. No.: 9,37e-122 Length: 1713  
 Score: 1311.50 Matches: 244  
 Percent Similarity: 68.0% Conservative: 76  
 Best Local Similarity: 51.91% Mismatches: 131  
 Query Match: 52.33% Indels: 19  
 DB: 4 Gaps: 4

US-10-004-219B-14 (1-455) x US-09-343-623-5 (1-1713)

```

QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 79 AAACCTGGTCTGCTACTTCCACCACTGGGCCAGTACAGACAGGAGGAGGCTCGCTTCCTG 138
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
DB 139 CCCAAGACTTGGACCCACGCTTTGACCCACCTCATCTACGCCCTTGGCTGGCATGCC 198
QY 42 AsnAsnGluIleThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 199 AACCAACCACTGAGCACCACCTGAGTGAATGACGAGACTCTCTACCAAGAGTTCAATGGC 258
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhedg 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTCGGC 318
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnAtgGlnThrPheIleThrSer 101
DB 319 ACTCAGAAGTTTACAGATATGTTAGCCAGCCGCAACACCTGTCAGACCTTTGTCAACTCG 378
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTCTCGCAAAATACAGCTTTGAGCGCTTGACCTTGACCTGGAGTACCCA 438
QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 439 GGAAGCCAGGGGAGGCGCTGCGGTAGACAGGAGCGCTTCACAAACCTGGTACAGACCTG 498
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
DB 499 GCCAATGCTTCCAGCAGGAGCCAGACCTCAGGGAAGGAGCGCTTCTCTGAGTGCA 558
QY 162 AlaValAlaIleGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLysSerGln 181
DB 559 GCGGTTCAGCTGGGACAGACCTATGCTGATGCTGATGAGTGGACGAGGTGACAAATGCCCC 618
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 619 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTTGGGAGAAGGTCAG 678
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 679 GGACATACAGCCCTTACAGAGGCAAGAGAGTGTGTCAGCAGCCAGCCCTCAAC 738
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
DB 739 GTGGATGCTGCTGCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATCCT 798
QY 242 GlyPheProThrTyrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyIleGly 261
DB 799 GGCATGCTTACTACGAGCGCTCTCTTCACTGGGCTCTCATCAGACCAACAGAGTGGGG 858
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
DB 859 GCCCAGCCAGCAGGCTTGGGACTCAGGCCCTTACCAAGGAAGGAGGAGTGTGGCC 918
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
DB 919 TACTATGAAGTCTGCTCTCTGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAG 972
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
DB 973 AAGGTGCCCTACATCTCTGGGACAAACAGTGGTGGGCTTTGATGATGTGGAGAGCTTC 1032

```

```

QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
DB 1033 AAAACCAAGGTCAGCTATCTGAGCAGAGAGGACTGGGGGGGGCCATGGTCTGGGCACTG 1092
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
DB 1093 GACTTAGATGACTTTGGCCGGCTTCTCTGCAACACGAGGCGCATACCCCTCATCCAGAGC 1152
QY 362 LeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnPro-IleGln 381
DB 1153 CTACGGCAGGAACGAATGGGTAAAGCCTTAACTGCCTGTCAATGTGAGGCCAGGTGT 1212
QY 381 uProlleThr-AlaAla-----ProSerGlySerGlyAsn- 392
DB 1213 GCCTGTGCACATGTGCTTCAGCTGTAGTCTTCCATATCTTGCCTTCAGGCCACCCAGAGC 1272
QY 393 -----GlySerGlySerSerSerSerGlyGlySerSerGlyGlySerG 407
DB 1273 TTGAAGTTTCCAAAACCAAGGTCAGGCTCTGAACCTGAGCATGGCCCAAGCCCTGGACAAG 1332
QY 407 ly---PheCysAlaValArgAlaAsnGlyLeuTyrProValAlaAsnAsnArgAsnAlaP 426
DB 1333 ACAGCTTCTGCCAGGGCAAGCTGATGGGCTCTATCCCAATCCTCGGGAACGGTCCAGCT 1392
QY 426 heTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPheA 446
DB 1393 TCTACAGCTGTGCAGCGGGGGGGCTGTTCAGCAAGAGTCCCGACAGCCCTGGTGTTC 1452
QY 446 spThrSerCysAspCysCysAsnTrp 454
DB 1453 GCAACTCTCTGCAATGTCTGCACCTGG 1478

```

## RESULT 13

```

US-08-694-915-1
; Sequence 1, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1433 base pairs

```

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-694-915-1

## Alignment Scores:

Pred. No.: 9,886-88 Length: 1433  
Score: 970.50 Matches: 182  
Percent Similarity: 69.11% Conservative: 73  
Best Local Similarity: 49.32% Mismatches: 105  
Query Match: 38.73% Indels: 9  
DB: 1 Gaps: 5

US-10-004-219b-14 (1-455) x US-08-694-915-1 (1-1433)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
DB 115 TACAACTGGTTGGCTTACCTTTACCACTGGTCCAGGACCGGAGAACCCAGGAAATTC 174  
QY 21 MetProAspAsnIleAspProCysLeuLeuThrHisLeuIleTyrAlaPheAlaGlyArg 40  
DB 175 ACCCTGAGATATGACCCCTTCCTATGCTCTCATCTCATCTATTCATTCATTCGCCAGCATC 234  
QY 41 GlnAsnAsnGluLeuThrThrLeuGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60  
DB 235 GAAACCAACAAAGGTATCATCAAGGACCAAGAGTGAAGTGTCTTACCAGACCATCAAC 294  
QY 61 GlyLeuLeuAsnLysAsnSerGlnLeuLeuThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
DB 295 AGTCTCAAAACCAAGAAATCCCAACTGAAATTCCTCTGTCCATTGGAGGGTACCTGTTT 354  
QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
DB 355 GGTTCACAAAGGTTCCACCTATGGTGGATTCTTCTACATCAGCTTGGAAATTCATTAC 414  
QY 101 SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
DB 415 TCCATAATCTGTTCTTGAGGAACCAATACTTTGATGGAGTGGATGTAAGTGGATCTAC 474  
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
DB 475 CCAGATCAGAAA-----GAAACACTCAT----TTCACTGTGCTGTATCATGAG 519  
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
DB 520 TTAGCAGAGACCTTTCAGAGGACTTCACAAATCCACCAGGAAAGCTTCTCTTGACT 579  
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
DB 580 GCGGGGTATCTGCAGGGAGGCAAAATGATTGATAACAGCTATCAAGTTCAGAAACTGGCA 639  
QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGly--- 199  
DB 640 AAGATCTGGATTTCATCAACTCCCTGTCTTGACTTCATGGGTCTTGGGAAAGGCC 699  
QY 200 ---TyrThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAla 218  
DB 700 CTTATCACTGGCCACACAGCCCTCTGACGAGGGGTGGCAGACAGAGGGCCAGCTCC 759  
QY 219 TyrLeuAsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLys 238  
DB 760 TACTACATGTGGAATGATGTGGGGTACTGGATACATAAGGGGAATGCCATCAGAGAAG 819  
QY 239 LeuIleValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThr 258  
DB 820 GTGGTCATGGGCATCCCATATGGGCACTCTTTCACACTGGCCCTCT---CGAGAAACC 876  
QY 259 GlyIleGlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGly 278

DB 877 ACCGTGGGGGCCCTCTGCTGAGCTGCTGGACCCATCACAGAGTCTTCAGGC 936  
QY 279 IleTrpAlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAsp 298  
DB 937 TTCCTGGCCCTATTATGAGATCTGCCAGTTCTCTGAAA---GGAGCCAAAGATCACGGGCTC 993  
QY 299 AlaProGlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIle 318  
DB 994 CAGGATCAGAGGTTCCCTACGCACTCAAGGGAACCACTGGTGGCTATGATGATG 1053  
QY 319 LysSerPheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetVal 338  
DB 1054 AAGAGTATGGAGACCAAGCTTCAGTTCTTAAAGATTTAAACCTGGAGAGCCATGATC 1113  
QY 339 TrpAlaIleAspLeuAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeu 358  
DB 1114 TGGTCTATTGACATGGATGACTTCACTGCAAACTCTGCAACCCAGGGCCCTTACCTCTT 1173  
QY 359 IleSerThrLeuLysAlaLeuGly 367  
DB 1174 GTCCAAGCAGTCAAGAGACCTTGGC 1200

## RESULT 14

US-08-694-915-3  
Sequence 3, Application US/08694915  
Patent No. 5811535

## GENERAL INFORMATION:

APPLICANT: Adamou, Julie  
APPLICANT: Kirkpatrick, Robert  
APPLICANT: Rosenberg, Martin  
TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-2799

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fast-Seq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,915  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Han, William T  
REGISTRATION NUMBER: 34,344  
REFERENCE/DOCKET NUMBER: ATG50017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5219  
TELEFAX: 610-270-5090  
TELEX:

## INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1526 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-694-915-3

## Alignment Scores:

Pred. No.: 1.09e-87 Length: 1526  
Score: 970.50 Matches: 182  
Percent Similarity: 69.11% Conservative: 73  
Best Local Similarity: 49.32% Mismatches: 105  
Query Match: 38.73% Indels: 9  
DB: 1 Gaps: 5

US-10-004-219B-14 (1-455) x US-08-694-915-3 (1-1526)

Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
Db 208 TACAAACTGGTTGCTACTTTTACCAACTGGTCCAGGACCGCAGGAACCAAGAAATTC 267

Qy 21 MetProAspAsnLeuAspProCysLeuCysThrHisLeuLeuTyrAlaPheAlaGlyArg 40  
Db 268 ACCCTGAGATATGACCCCTTCTATGCTCTCATCTCATCTATCATCTATGCTCCAGCATC 327

Qy 41 GlnAsnAsnGluLeuThrIleGluTrpAsnValThrLeuTyrGlnAlaPheAsn 60  
Db 328 GAAACCAACAAGTTATCATCAAGGACCAAGAGTGAAGTGATGCTCTACCAAGACCAATCAAC 387

Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
Db 388 AGTCTCAAAACCAAGAAATCCCAAACTGAAATTCCTTTGCTCCATTGGAGGGTACCTGTT 447

Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
Db 448 GTTCCAAAGGTTCCACCTATGTTGGATCTTCTACATCACGCTTGGAAATTCATTAAC 507

Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
Db 508 TCCATAATCCTGTTCTCAGGAACCAATAACTTTGATGACTGGATGAAGTGGATCTAC 567

Qy 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
Db 568 CCAGATCAGAAA-----GAAACACTCAT---TTCACTGTGCTGATCATGAG 612

Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
Db 613 TTAGCAGAAGCCTTTCAGAGGACTTCACAAATCCACAGGAAGGCTTCTGTGACT 672

Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
Db 673 GCGGCGGTATCTGCAGGAGGCAAAATGATTGATAACAGCTATCAAGTTGAGAACTGGCA 732

Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGly--- 199  
Db 733 AAAGATCTGGATTTTCATCAACCTCTCTCTCTGCTTCCATGCTTGGGAAAAGCCC 792

Qy 200 ---TyrThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAla 218  
Db 793 CTTATCATGCGCCACACAGCCCTTGTAGCAAGGGTGGCAGGACAGAGGGCCAGCTCC 852

Qy 219 TyrLeuAsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyValaProAlaGluLys 238  
Db 853 TACTACATGTGGAATATGCTGTGGGTACTGGATACATAAGGGAATGCCATCAGAGAAG 912

Qy 239 LeuIleValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerArgThr 258  
Db 913 GTGGTCATGGGCATCCCATATATGGCACTCTTTCACACTGGCCTCT---GCAGAAACC 969

Qy 259 GlyIleGlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGly 278  
Db 970 ACCGTGGGGGGCCCTCTGCGCTGAGCTGTGGACCATCAGAGAGCTTCAGGC 1029

Qy 279 IleTrpAlaTyrTyrGluIleCysThrPheLeuLysAsnGlyValaThrGlnGlyTrpAsp 298  
Db 1030 TTCCTGGCCTATTATGAGATCTGCCAGTCTCTGAAA---GGAGCCAGATCAGCGGCTC 1086

Qy 299 AlaProGlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIle 318  
Db 1087 CAGGATCAGCAGGTTCCCTACGCAGTCAAGGGGAACCAAGTGGTGGGTGCTATGATGATG 1146

Qy 319 LysSerPheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetVal 338  
Db 1147 AAGAGTATGGAGACCAAGTTTCAGTTCTTAAGAAATTAACCTGGGAGGACCATGATC 1206

Qy 339 TrpAlaIleAspLeuAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeu 358  
Db 1207 TGGTCTATTGACATGGATGACTTCACTGGCAAAATCTCTCAACACGAGGCCCTTACCTCTT 1266

Qy 359 IleSerThrLeuLysLysAlaLeuGly 367  
Db 1267 GTCCAAGCAGTCAAGAGAAGCCTTGGC 1293

## RESULT 15

PCT-US94-07754-4  
; Sequence 4, Application PC/TUS9407754  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR  
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SPENSLER HORN JUBAS & LUBITZ  
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 90067

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/07754  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOWELLS, STACY L.  
; REGISTRATION NUMBER: 34,842  
; REFERENCE/DOCKET NUMBER: FD 3665  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/455-5100  
; TELEFAX: 619/455-5110  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1681 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; IMMEDIATE SOURCE:  
; CLONE: YKL-40  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 135..1681

Alignment Scores:  
Pred. No.: 7.99e-87 Length: 1681  
Score: 962.50 Matches: 194  
Percent Similarity: 65.82% Conservative: 64  
Best Local Similarity: 49.49% Mismatches: 122  
Query Match: 38.41% Indels: 13  
DB: 5 Gaps: 6

US-10-004-219B-14 (1-455) x PCT-US94-07754-4 (1-1681)

Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
Db 135 TACAAACTGGTCTGCTACTACTACACAGCTGGTCCCAAGTACCGGAGGATGGGAGCTGC 194

Qy 21 MetProAspAsnLeuAspProCysLeuCysThrHisLeuLeuTyrAlaPheAlaGlyArg 40



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: June 29, 2003, 23:54:39 ; Search time 165.913 Seconds  
(without alignments)  
4070.941 Million cell updates/sec

Title: US-10-004-219B-14  
Perfect score: 2506  
Sequence: 1 YQLTCYFTNWAQYRGLGRF.....QQNCOAGLVFDTSCDCNWA 455

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO\_spool/US10004219/runat\_24062003\_160230\_8285/app\_query.fasta\_1.1877  
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnbp -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10004219 @CGN 1.1 123 @runat\_24062003\_160230\_8285  
-NCPUE=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2506	100.0	1625	9 US-10-004-219B-2
2	2089.5	83.4	1525	9 US-10-004-219B-3
3	1251	49.9	1038	9 US-10-202-436A-14
4	985.5	39.3	1925	9 US-10-097-340-44

5	957.5	38.2	1681	10 US-09-215-077A-4
6	957.5	38.2	1681	10 US-09-262-213A-4
7	923.5	36.9	1391	10 US-09-822-830A-402
8	873.5	34.9	1474	10 US-09-765-231A-8
9	711.5	28.4	1470	9 US-10-218-743-40
c 10	711.5	28.4	1470	9 US-10-218-743-42
c 11	711.5	28.4	1527	9 US-10-218-743-37
c 12	711.5	28.4	1527	9 US-10-218-743-39
c 13	711.5	28.4	1621	9 US-10-218-743-34
c 14	711.5	28.4	1621	9 US-10-218-743-36
c 15	708.5	28.3	1608	9 US-10-218-743-20
c 16	708.5	28.3	1608	9 US-10-218-743-22
c 17	708.5	28.3	1665	9 US-10-218-743-17
c 18	708.5	28.3	1665	9 US-10-218-743-19
c 19	708.5	28.3	1752	9 US-10-218-743-14
c 20	708.5	28.3	1752	9 US-10-218-743-16
c 21	646	25.8	449	10 US-09-960-352-678
c 22	472	18.8	1271	9 US-09-923-844B-1
c 23	444.5	17.7	1140	9 US-09-938-842A-1684
c 24	444	17.7	1797	10 US-09-974-300-646
c 25	380.5	15.2	410	10 US-09-960-352-7364
c 26	380.5	15.2	414	10 US-09-960-352-13050
c 27	380.5	15.2	415	10 US-09-960-352-2589
c 28	380.5	15.2	424	10 US-09-960-352-3057
c 29	378.5	15.1	418	10 US-09-960-352-3072
c 30	363	14.5	399	10 US-09-960-352-265
c 31	361.5	14.4	411	10 US-09-960-352-14919
c 32	352	14.0	398	10 US-09-960-352-14460
c 33	347.5	13.9	384	10 US-09-960-352-5721
c 34	342	13.6	427	10 US-09-960-352-3315
c 35	341	13.6	438	10 US-09-960-352-1882
c 36	339	13.5	422	10 US-09-960-352-5957
c 37	339	13.5	423	10 US-09-960-352-12100
c 38	338	13.5	430	10 US-09-960-352-1189
c 39	336	13.4	419	10 US-09-960-352-10172
c 40	336	13.4	427	10 US-09-960-352-4978
c 41	335	13.4	426	10 US-09-960-352-4979
c 42	334	13.3	419	10 US-09-960-352-4053
c 43	334	13.3	422	10 US-09-960-352-10642
c 44	332	13.2	410	10 US-09-960-352-8317
c 45	331	13.2	429	10 US-09-960-352-14230

ALIGNMENTS

RESULT 1  
US-10-004-219B-2  
; Sequence 2, Application US/10004219B  
; Publication No. US20030087414A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1625  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human  
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid  
; OTHER INFORMATION: sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase  
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence  
; FEATURE:

NAME/KEY: CDS  
LOCATION: (104)...(1531)  
US-10-004-219B-2

## Alignment Scores:

Pred. No.: 3,546-280 Length: 1625  
Score: 2506.00 Matches: 455  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-004-219B-14 (1-455) x US-10-004-219B-2 (1-1625)

QY	1	TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe	20
DB	167	TACGAGCTGACATGCTTACCACTAGTGGCCAGTACCGCCAGGCTGGGGCGGCTTC	226
QY	21	MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg	40
DB	227	ATGCTGTGACAACTGACCCCTGCTCTGTACCACCTGATCTACGCTTTGCTGGAGG	286
QY	41	GlnAsnAsnGluIleThrThrIleGluTyrAsnAspValThrLeuTyrGlnAlaPheAsn	60
DB	287	CAGAACACGAGATCAACCATCGAATGGAACGATGCTCTTACCAGGCTTTCAAT	346
QY	61	GlyLeuLeuAsnIleAsnSerGlnLeuLeuThrLeuAlaIleGlyGlyTyrAsnPhe	80
DB	347	GGCTGTAAATAAAGAACAGCCAGCTGAAACTCTCTGGCCATTGGAGCTGGAATTC	406
QY	81	GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr	100
DB	407	GGGACTGCCCTTTCACTGCCATGGTTTCTACTCTTGAGAACCGCCAGACTTTTCATCAC	466
QY	101	SerValIleIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr	120
DB	467	TCAGTCATCAAAATCTCGGCCAGTAGATGTTGACGGCTGGAGCTTTGACTGGGAGTAC	526
QY	121	ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu	140
DB	527	CTGGCTCTCGTGGAGGCTCTCTCAGGACAGCATCTCTTCACTGCTCTGGTGAGGAA	586
QY	141	MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr	160
DB	587	ATGCGTGAAGCTTTTGAGCAGGAGGCGCAAGCAGATCAACAGCCAGGCTGATGCTCACT	646
QY	161	AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer	180
DB	647	GCTGCAGTAGCTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCAACTGTCA	706
QY	181	GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr	200
DB	707	CAGTACTCGACTACATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT	766
QY	201	ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu	220
DB	767	ACTGGAGAGACAGCCCTCTACAAATACCCGACTGACACCGCAGCAGCAACGCTACCTC	826
QY	221	AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle	240
DB	827	AATGTGGATTATGTCATGAACCTACTGGAAGGACAAATGGAGCAGCAGCTGAGAAGCTCATC	886
QY	241	ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle	260
DB	887	GTTGGATTCTTACCTATGACACAACTTCACTTCTGAGCAACCCCTCCCAACTGGAAT	946
QY	261	GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp	280
DB	947	GGTGGCCCTCCTCTGGTCTGCTGCTGGCCCTATGCCAAGAGCTGGATCTGG	1006
QY	281	AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro	300
DB	1007	GCTTACTACGAGATCTGTACTCTCTGAAAAATGGAGCCACTCAGGATGGGATGCCCT	1066

QY	301	GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer	320
DB	1067	CAGGAAGTGCCTTATGCTTATCAGGCGCAATGTGTGGTGGCTATGACACATCAAGAGC	1126
QY	321	PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla	340
DB	1127	TTTCGATATTAAAGCTCAATGGCTTAAGCACAACAATTTGGAGGCGCCATGCTCTGGGCG	1186
QY	341	IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer	360
DB	1187	ATTGATCTGGATGACTTCACTGGCCTTTCTTCAACACAGGCGCAAGTTTCCCTAATCTCC	1246
QY	361	ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle	380
DB	1247	ACCTGAGAAGAGCCCTCGGCTTCAGAGTGCAGTGCAGGTTCACGGCTCCAGCTCAGCCCAT	1306
QY	381	GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly	400
DB	1307	GAGCCAATAACTGCTGCTCCAGTGGCAGCGGGAACGGGAGGAGTAGCAGCTCTGGA	1366
QY	401	GlySerSerGlyGlySerGlyPheCysAlaValAlaArgAlaAsnGlyLeuTyrProValAla	420
DB	1367	GGCAGCTCGGAGGAGTGGATTCTGTGCTGTGAGAGCAACGGCTCTTACCCCGTGGA	1426
QY	421	AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnAsnCysGln	440
DB	1427	ATAACAGAAATGCTTCTGGCAGTGGTGAATGATGATGATGATGATGATGATGATGAT	1486
QY	441	AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrpAla	455
DB	1487	GGCGGGCTGTCTTCGACACAGCTGTGATTGCTGCAACTGGGCA	1531

## RESULT 2

US-10-004-219B-3  
; Sequence 3, Application US/10004219B  
; Publication No. US2003008741A1  
; GENERAL INFORMATION:  
; APPLICANT: Macrozyme  
; APPLICANT: Aerts, Johannes M.F.G.  
; APPLICANT: Boot, Rolf G.  
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and  
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in  
; TITLE OF INVENTION: which mucus is involved or infection diseases  
; FILE REFERENCE: 2183-5136US  
; CURRENT APPLICATION NUMBER: US/10/004,219B  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3

LENGTH: 1525  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Combined DNA/RNA Molecule: mouse  
OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid  
OTHER INFORMATION: sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase  
OTHER INFORMATION: cDNA sequence and deduced amino acid sequence  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1419)  
US-10-004-219B-3

## Alignment Scores:

Pred. No.: 5,796-232 Length: 1525  
Score: 2089.50 Matches: 369  
Percent Similarity: 90.05% Conservative: 40  
Best Local Similarity: 81.28% Mismatches: 42  
Query Match: 83.38% Indels: 3  
DB: 9 Gaps: 1



```
Db 242 TACTACGAGATCTGTACCTCTCTGAAAAATGGAGCCACTCAGGATGGGATGCCCTCAG 301
Qy 302 GluValProTyrAlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleLysSerPhe 321
Db 302 GAAGTGCCTTATGCCCTATCAGGCAATGTGTGGTTGGCTATGACACGTCAGAGCTTC 361
Qy 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
Db 362 GATATTAAAGGCTCAATGGCTTAAAGCACACAAATCTGGAGGCGCCATGTGTGGGCCATT 421
Qy 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
Db 422 GATCTCGATGACTCTACCTGCACTTTCTGCAACAGGCGCAAGTTTCCCTAATCTCCACC 481
Qy 362 LeuLysLysAlaLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGlu 381
Db 482 CTGAAGAAGGCCCTCGGGCTGCAGATGCAGATGCAGGCTCCAGCTCAGCCCATITGAG 541
Qy 382 ProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGlyGly 401
Db 542 CCAATAACTGTCTCCAGTGGCAGCGGGAACGGGAGCGGAGTAGCAGCTCTGGAGGC 601
Qy 402 SerSerGlyGlySerGlyPhe---CysAlaValArgAlaAsnGlyLeuTyrProValAla 420
Db 602 AGCTCGGAGGCGAGTGGATTTCTGTGCTGGCAGAGCAAAACGAGCTCTAAACCCCGTGGC 661
Qy 421 Asn---AsnArgAsnAlaPheTyrHisCysValAsnGlyValThrTyrGln-GlnAsn- 438
Db 662 AATTACCAAGAGATGCCCTTCTGGGCACTGCGTGAATGAGTGCACGTACAGGACAGACT 721
Qy 439 CysGlnAlaGlyLeuVal-PheAspThrSerCysAspCysCysAsnTrpAla 455
Db 722 TGCCAGCGCGGGCTGTCTCTTCGAGACAGCTGTGTAATGCTGCAACTGGGCA 773

RESULT 4
US-10-097-340-44
; Sequence 44, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNANARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
```

```
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-44

Alignment Scores:
Pred. No.: 6,74e-104 Length: 1925
Score: 985.50 Matches: 195
Percent Similarity: 65.74% Conservative: 64
Best Local Similarity: 49.44% Mismatches: 120
Query Match: 39.33% Indels: 15
DB: 7

US-10-004-219B-14 (1-455) x US-10-097-340-44 (1-1925)

Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Db 190 TACAACCTGGTCTGCTACTTACACAGCTGGTCCAGTACCCGGAAGCGATGGAGCTGC 249
Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Db 250 TTCCAGATGCCCTTGACCGCTTCTCTGTATCCACATCATCATGCTTTGCCAATATA 309
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
Db 310 AGCAACGATCATCGACACCTGGGAGTGAATGATGACGCTCTACGGCATGCTCAAC 369
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyTrpAsnPhe 80
Db 370 ACACCTCAAGAACAGGACCCCAACCTGAAGACTCTCTGTCTGTGCGAGGATGAACTTT 429
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Db 430 GGGTCTCAAGATTTTCCAAGATAGCTCTCAACACCCAGAGTCGCGGACTTTTCATCAAG 489
Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Db 490 TCAGTACCGCCCATCTCTGGCCACCCATGCTTTGATGGCTGGACCTTGCCTGGCTCTAC 549
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 550 CTGGAGCGGAGA-----GACAAACAGCATTTTACCACCCTAATCAAGGAA 594
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 595 ATGAAGCGCGAATTTATAAAGGAAGCC---CAGCCAGGGAAGAAAGCAGCTCTGCTCAGC 651
Qy 161 AlaAlaValAlaAlaGlyLysSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Db 652 GCAGCACTGTCTGCGGGGAAGGTCACTACCATGACAGCTATGACATTCGCAAGATATCC 711
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 712 CAACACCTGGATTTTCATTAGCATCATGACCTACGATTTTCATGGAGCTGGCGTGGGACC 771
Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 772 ACAGGCCATCATCAGTCCCTCTGTCGAGGTCAAGGATGCAAGTCCCTGACAGATTCAGC 831
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
Db 832 AACACTGACTATGCTGTGGGTGATGTTGAGGCTGGGGGCTCTCCGCCAGTAAAGTGGTG 891
Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Db 892 ATGGGCATCCCACTCTCGGAGGAGGCTTCACTCTGGCTTCT---TCTGAGACTGGTGT 948
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrProTyrAlaLysGluSerGlyIleTrp 280
```



949	GGAGCCCCATCTCAGGACCGGGAATTCAGGCCGGTTCACCAAGGAGGAGGACGGACCCCTT	1008
281	AlaTyrTyrGluIleCysThrPheLeuLeuYAsnGlyAlaThrGlnGlyTrpAspAlaPro	300
1009	GCCTACTATGAGATCTGTGATCTTCTCCGCG--GGAGCCACAGTCCATAGAACCCCTCGGC	1065
301	GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer	320
1066	CAGCAGTCCCTTATGTCACCAAGGGCAACAGTCGTGGTAGGATACGACACCAAGGAAAGC	1125
321	PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla	340
1126	GTCAAAAGCAAGGTGCAGTACCTCGAAGGATAGCAGCTGGCAGCGCCATGTGTATGGGCC	1185
341	IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGly--LysPheProLeuIle	359
1186	CTGACCTTGGATGACTTCCAGGGCTCTTCTGGCGGCAGGATCTGGCGCTTCCCTCTCACC	1245
360	SerThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSer-----CysThrAla	375
1246	AATGCCATCAAGGATGCACCTCGCTGCAACGTAGCCCTCTGTCTTCTGCACACAGCAGCGGGG	1305
376	ProAlaGlnProIleGluProIleThrAlaAlaProSerGly	389
1306	CCAAGATGCCCGTCCCTCTCG-----GCTCCAGCTGGC	1341

255 AGCAACGATCATCATCGACACCTGGGAGTGGAAATGATGTGACGCTCTACGGCATGCTCAAC 314  
 61 GlyLeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80  
 315 ACACCTCAACAACACGACCAACCCCAACCTGAAGACTCTCTGTCTGCGAGGATGGAACTTT 374  
 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
 375 GGGTCTCAAAAGATTTTCCAAAGATAGGCTCCAAACACCCAGAGTAGTCCGGACTTTTCATCAAG 434  
 101 SerValIleIysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTyrGluTyr 120  
 435 TCAGTACCGGCATTTCTCGCACCACTGGCTTTGATGGCGTGACCTTGCTGGCTCTAC 494  
 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
 495 CCTGGAGCGGAGA-----GACAAACACCATTTTACCACCCCTAATCAAGGAA 539  
 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
 540 ATGAAGCCCGAATTTATAAAGGAAGCC---CAGCCAGGGAAAAAGCAGCTCTGCTCAGC 596  
 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
 597 GCAGCACTGTCTCGGGGAAGGTACCATTTGACAGCAGCTATGACATTCGACAGATATCC 656  
 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyr 200  
 657 CAACACCTGGATTTCATTAGCATCATGACCTACGATTTTCATGCGCGCTGGCGTGGGACC 716  
 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
 717 ACAGGCGCATCAAGTCCCTCCAGCGAGGTCAGGAGGATCAAGATCTCGTACATTCAGC 776  
 221 AsnValAspTyrValMetAsnTyrTyrLysAspAsnGlyValaProAlaGluLysLeuLe 240  
 777 AACACTGACTATGCTGTGGGTACATGTTGAGGCTGGGGGCTCTGCGACATGAGCTGGT 836  
 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260  
 837 ATGGGCATCCCACTTCGGGAGGAGCTTCACCTGGCTTCT---TCTGAGACTGGTGTT 893  
 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyr 280  
 894 CCAGCGCCAATCTCAGGACCGGGAATTCAGGCGGTTCCACCAAGGAGGAGGACGCTT 953  
 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyValaThrGlnGlyTyrAspAlaPro 300  
 954 GCCTACTATGAGATCTGTGACTTCCTCCGCGC---GGAGCCACAGTCCATAGAACCTCGGC 1010  
 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleLysSer 320  
 1011 CAGCAGTCCCTTATGCCACCAAGGGCAACCACTGGGTAGATACGACGACGAGGAAAGC 1070  
 321 PheAspIleLysAlaGlnTyrLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAla 340  
 1071 GTCAAAAGCAAGGTGCAGTACCTGAAGGATAGGAGCTGGCGAGCGGCATGATGGGCC 1130  
 341 IleAspLeuAspPheThrGlyThrPheCysAsnGlnGly---LysPheProLeuLe 359  
 1131 CTGGACCTGGATGACTTCAGGGGCTCCTCTGGCGGCGAGGATTCGCGCTTCCCTCTCACC 1190  
 360 SerThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnPro 379  
 1191 AATGCCATCAAGGATGCATCTCGTGC-AACGTAGCCCTCTGTTCTGTCGACACGACGACGGG 1249  
 380 IleGluProIleThrAlaAlaProSerGlySerGly 391  
 1250 GCCAAGGAT-----GCCCGCTCCCTCTGCGC 1276

RESULT 6  
US-09-262

Sequence 4, Application US/09262213A  
Patent No. US20020090658A1

## GENERAL INFORMATION:

APPLICANT: PRICE, PAUL  
APPLICANT: JOHANSEN, JULIA  
TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF  
TITLE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES  
FILE REFERENCE: 407T-895412US  
CURRENT APPLICATION NUMBER: US/09/262,213A  
CURRENT FILING DATE: 1999-03-04  
PRIOR APPLICATION NUMBER: US 08/089,989  
PRIOR FILING DATE: 1993-07-09  
PRIOR APPLICATION NUMBER: PCT/US94/07754  
PRIOR FILING DATE: 1994-07-08  
PRIOR APPLICATION NUMBER: US 08/581,527  
PRIOR FILING DATE: 1996-04-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4

LENGTH: 1681

TYPE: DNA

ORGANISM: Homo sapiens

US-09-262-213A-4

## Alignment Scores:

Pred. No.:	9,61e-101	Length:	1681
Score:	957.50	Matches:	194
Percent Similarity:	65.56%	Conservative:	63
Best Local Similarity:	49.49%	Mismatches:	123
Query Match:	38.21%	Indels:	13
DB:	10	Gaps:	6

US-10-004-219B-14 (1-455) x US-09-262-213A-4 (1-1681)

QY	1	TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe	20
DB	135	TACAAACTGGTCTGCTACTACTACCAAGCTGGTCCAGTACCGGGAAGCGATGGAGCTGC	194
QY	21	MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg	40
DB	195	TTCCAGATGCCCTTGACCGCTCTCTGTGTACCAATCATCTACAGCTTGGCAATATA	254
QY	41	GlnAspAspGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn	60
DB	255	AGCAACGATCATCATCGACACCTGGGAGTGGAAATGATGACGCTCTACGCGATGCTCAAC	314
QY	61	GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe	80
DB	315	ACACTCAACAACAACGAACCCCAACCTGAAGACTCTCTGTGTGCGGAGATGGAACTTT	374
QY	81	GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr	100
DB	375	GGGTCTCAAGATTTTCCAGATAGCTTCAACACCCAGAGTCCCGGACTTTCATCAAG	434
QY	101	SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr	120
DB	435	TCAGTACCGCCATTTCTGCGCACCCATGCTTTGATGGCGGTGACCTTGCCTGCTCTAC	494
QY	121	ProGlySerArgGlySerProProGlnAspLysPheThrValLeuValGlnGlu	140
DB	495	CCTGGACGGAGA-----GACAAACACCATTTTACCAACCCCTAATCAAGAA	539
QY	141	MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr	160
DB	540	ATGAAGGCCGAAATTTAAGGAGGCC---CAGCCAGGGAAGAGAGCTCTGCTCAGC	596
QY	161	AlaAlaValAlaIleGlyLeuSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer	180
DB	597	GCAGCACTGTCTCGCGGAGGTCACCATTCAGCAGCAGCTATGACATGCGCAAGATATCC	656
QY	181	GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr	200
DB	657	CAACACCTGGATTTTACATGATCATGACCTACGATTTTTCATGGCGCCTGGGAGCC	716

QY	201	ThrGlyGluAsnSerProLeuTyrIleTyrTrpThrAspThrGlySerAsnAlaTyrLeu	220
DB	717	ACAGGCCATCACAGTCCCTCAGCGGAGGTCAAGTCTCTGACACGATTCAGC	776
QY	221	AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeu	240
DB	777	AACACTGACTATGCTGGGTGACATGTTGAGGCTGGGGCTCTCTGCCAGTAGTGGT	836
QY	241	ValGlyPheProThrTyrHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle	260
DB	837	ATGGGCATCCACCTTCGGAGGAGCTTCACTCTGGCTTCT---TCTGAGACTGGT	893
QY	261	GlyAlaProThrSerGlyValAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIle	280
DB	894	CCAGCGCAATCTCAGGACCGGGAATTCAGCGCGTTTCAACCAAGGAGGAGCCCTT	953
QY	281	AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAla	300
DB	954	GCCTACTATGAGATCTGTGACTTCTCTCCGC---GGAGCCACAGTCCATAGAAC	1010
QY	301	GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLys	320
DB	1011	CAGCAGGTCCTTATGCCACCAAGGCAACAGTGGTAGGATACACACGACGAGAA	1070
QY	321	PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrp	340
DB	1071	GTCAAAAGCAAGTGCAGTACCTGAAGGATAGGACGAGTGGCAGCGCCATGTAT	1130
QY	341	IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGly---LysPheProLeu	359
DB	1131	CTGGACCTGGATGACTTCCAGGCTCTCTTCTGCGCCAGGATCTCGCTTCCCT	1190
QY	360	SerThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAla	379
DB	1191	ATGCGCATCAAGATGACCTGCTGC---ACGTAGCCCTCTGTTCTGTCACACAG	1249
QY	380	IleGlyProIleThrAlaAlaProSerGlySerGly	391
DB	1250	GCCAAGGAT-----GCCCGTCCCGTCTGCGC	1276

## RESULT 7

US-09-822-830A-402  
Sequence 402, Application US/09822830A  
Patent No. US20020142952A1  
GENERAL INFORMATION:  
APPLICANT: Genetics Institute, Inc.  
APPLICANT: Wong, Gordon G.  
APPLICANT: Clark, Hilary  
APPLICANT: Fectel, Kim  
APPLICANT: Agostino, Michael J.  
APPLICANT: Howes, Steven H.  
APPLICANT: Resnick, Richard J.  
APPLICANT: Gulukota, Kamalak  
APPLICANT: Graham, James R.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
FILE REFERENCE: GIN 6402  
CURRENT APPLICATION NUMBER: US/09/822,830A  
CURRENT FILING DATE: 2001-03-29  
PRIOR APPLICATION NUMBER: 60/195,604  
PRIOR FILING DATE: 2000-04-06  
NUMBER OF SEQ ID NOS: 631  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 402  
LENGTH: 1391  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1353  
OTHER INFORMATION: n-a, c, g, or t  
US-09-822-830A-402

## Alignment Scores:

Pred. No.: 6,246-97 Length: 1391  
Score: 923.50 Matches: 176  
Percent Similarity: 67.48% Conservative: 73  
Best Local Similarity: 47.70% Mismatches: 111  
Query Match: 36.85% Indels: 9  
DB: 10 Gaps: 5

US-10-004-219B-14 (1-455) x US-09-822-830A-402 (1-1391)

```
QY 1 TyrGlnLeuThrCysTyPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 76 TACAAACTGGTTGGGACTTTTACCAACTGGTCCCAGGACCCGAGAACCCAGGAAATTC 135

QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
DB 136 ACCCTGAGATATGACCCCTTCTATGCTCTCATCTCATCTATTCATTCGCCAGCATC 195

QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
DB 196 GAAACCAACAAGGTTATCATCAAGGACAAGAGTGAAGTGATGCTCTACCAAGCATCAAC 255

QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 256 AGTCTCAAAACCAAGAATCCCAAACTGAAATTCCTTGTCCATTGGAGGGTACCTGTT 315

QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
DB 316 GGTTCAAAGGTTTCACCTATGTTGGATTCTTCTACATCAGCTTGGAAATTCATTAC 375

QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
DB 376 TCATAATCTCTTCTGAGGAACCACTACTTTGATGGAGTGAAGTGAAGTGGATCTAC 435

QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
DB 436 CCAGATCAGAAA-----GAAACACTCAT---TTCACCTGTGCTGATTCATGAG 480

QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
DB 481 TTAGCAGAAGCTTTTCAGAGGACTTCACAAATCCACCAGGAAGGCTCTCTTGACT 540

QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
DB 541 GCGGGCGTATCTGAGGAGGCAATGATTGATAACAGCTATCAAGTTTGAGAACTGGCA 600

QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGly--- 199
DB 601 AAAGATCTGGATTTTCATCAACCTCTGCTTGTGACTTCCATGGGTCTTTGGGAAAGCCC 660

QY 200 ---TyrThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAla 218
DB 661 CTTATCATTGGCCACACAGACCTCTGAGCAAGGGTGGCAGACAGAGGGCCAGCTCC 720

QY 219 TyrLeuAsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLys 238
DB 721 TACTACATGTGGAATATGCTGGGGTACTGGATACATAAGGGAATGCCATCAGAGAAG 780

QY 239 LeuIleValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThr 258
DB 781 GTGGTCATGGGATCCCAACATATGGGCACTCCTTCACATGGCCCTCT---GCAGAAACC 837

QY 259 GlyIleGlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGly 278
DB 838 ACCGTGGGGCCCTTGCCTCTGGCCCTGGAGCTGTGGACCCATCACAGATCTTCAGGC 897

QY 279 IleTrpAlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAsp 298
DB 898 TTCCTGGCCCTATTATGAGATCTGCCAGTTCCTGAAA---GGAGCCCAAGATCAGAGGCTC 954

QY 299 AlaProGlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIle 318
DB 955 CAGGATCAGAGGTTCCCTTACGCACTCAGGGGACCCCGTGGGTGGCTATGATGATGTG 1014
```

```
QY 319 LysSerPheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetVal 338
DB 1015 AAGAGTTTGGGGCCCAAGTTTCAGTTCTTAAGAATTTAAACCTGGGGGTGCTTATT 1074

QY 339 TrpAlaIleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeu 358
DB 1075 TGGTCTTTTGACATGGTTGACTTTCACCTGGCAATCTCTCAACCGGGGCCCTTCCCTCTT 1134

QY 359 IleSerThrLeuLysLysAlaLeuGly 367
DB 1135 GTCCAGCAGTCAGAGAGGCTTGGC 1161
```

## RESULT 8

US-09-765-231A-8  
; Sequence 8, Application US/09765231A  
; Patent No.: US20020119452A1  
; GENERAL INFORMATION:  
; APPLICANT: Searle/Monsanto  
; APPLICANT: Phippard, Deborah  
; APPLICANT: Vasanthakamur, Geetha  
; APPLICANT: Dotson, Stanton  
; APPLICANT: Ma, Xiao-Jun  
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,  
; FILE OF INVENTION: vectors, and cells  
; FILE REFERENCE: SO-3221 PR  
; CURRENT APPLICATION NUMBER: US/09/765,231A  
; CURRENT FILING DATE: 2001-01-18  
; NUMBER OF SEQ ID NOS: 82  
; SEQ ID NO 8  
; LENGTH: 1474  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-765-231A-8

## Alignment Scores:

Pred. No.: 4,246-91 Length: 1474  
Score: 873.50 Matches: 181  
Percent Similarity: 67.55% Conservative: 73  
Best Local Similarity: 48.14% Mismatches: 106  
Query Match: 34.86% Indels: 16  
DB: 10 Gaps: 5

US-10-004-219B-14 (1-455) x US-09-765-231A-8 (1-1474)

```
QY 1 TyrGlnLeuThrCysTyPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 137 TACAAACTGGTTGGTACTTTTACCAACTGGGTCCCAGGACCCGAGAACCCAGGAAATTC 196

QY 20 eMetProAsp-AsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPhe-AlaGly 39
DB 197 CACCCCTGAGGATATTCACCCCTTCTATGCTCTCATCTCATCTATTCATTCGCCAGC 256

QY 40 ArgGlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPhe 59
DB 257 ATCGAAAAACAAGGTTATCATCAAGGACAAGAGTGAAGTGATGCTCTTACCAAGCATC 316

QY 60 AsnGly-LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAs 79
DB 317 AACAGTTCTCAAAACCAAGAATCCCAAACTGAAATTCCTTGTCCATTGGAGGGTACCT 376

QY 79 nPheGlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIle 99
DB 377 GTTTGGTTCCAAAGGGTTCCACCTATGTTGGATTCTTCTACATCAGCTTGGAAATTCAT 436

QY 99 eTherSerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGl 119
DB 437 TAACTCCATAATCTCTGTTTCTGAGGAACCAATACTTTGATGGAGTGGATGTAAGCTGGAT 496

QY 119 uTyProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlu 139
DB 497 CTACCCAGATCAGAAA-----GAAACACTCAT---TTCACCTGTGCTGATTCATCA 541
```



1428	CGAATCGATGTTATGTTGGAAACATGGTCGGTTTATCATATAA---GTTGATCCATACACA	137	
Qy	22	ProAspAsnIleAspProCysLeuLeuIleThrHisLeuIleTyrAlaPheAlaGlyArgGln	41
Db	1371	ATTGAAGATATGATCTCTTCAAAATGACTCATTTGATGATGGTTTGGCTAAATCGAT	1312
Qy	42	AsnAsnGluIleThrThr-----leGluTrp	50
Db	1311	GAATACAAATACACCATTCAAAGTTTTTGATCCATTTCAAGATGATAACCATTAACCTCATGG	1252
Qy	51	AsnAspValThrLeuTyrGlnAlaPheAsnGlyLeuLeuAsnIleAsnSerGlnLeuLeu	70
Db	1251	---GAAAAACACGGGTATGAACCTTTCAACAACTTCAGATTTGAAGAATCCAGAAATGGACC	1195
Qy	71	ThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPheThrAlaMetValSer	90
Db	1194	ACCATGATTTTCATGGTGGTGGTATGAAGTTTCAGAAAATATTCCGGATATGGCAGCC	1135
Qy	91	ThrProGluAsnArgGlnThrPheIleThrSerValIleLeuLeuPheLeuArgGlnTyrGlu	110
Db	1134	AATCCAACATATCGTCAGCAATTTGTTCAATCAGTTTGGACTTTTTCGAAGAATACAAA	1075
Qy	111	PheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg---GlySerProProGln	129
Db	1074	TTCCGATGGCCATGATTTGGATTCGGGAATTCCTCGAATCACGGTTAGGCAATCCTAAAATC	1015
Qy	130	AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla	149
Db	1014	GATAAACAAACTATTAAACATTAGTTAGAGAACTTTAAGAGGCAATTTGAACCTTTCGGC	955
Qy	150	LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValaAlaGlyIleSerAsn	169
Db	954	-----TACTTGTTCAGCTCCGCGAGTATCACCCGGTAAAGATAAA	916
Qy	170	IleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet	189
Db	915	ATTGACGTAGCTTATGAGCTCAAGAAATTTGAACCAATTTGTTCCGATTTGGATGAATGTCATG	856
Qy	190	ThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLys	209
Db	855	ACTTATGATTACCATGGCGATGGGAATATGTTTCGGCCATTAATGCTCGTTGTATPAA	796
Qy	210	TyrProThrAspThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyr	228
Db	795	CGACCCGATGAACGGATGAATTTGCACACTTCTCAATGTCAACTACACCATGCACAT	736
Qy	229	TrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHis	248
Db	735	TATTTGAACAATGGCGCTACTCGAGACAACTTTGTTATGGGTGTCCATTCTATGTCGT	676
Qy	249	AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyValaProThrSerGlyAlaGly	268
Db	675	GCTTGAGACATCGAAGATCGAAGCAAAAGTCAAACCTTGGCGATCCGGCCAAAGCATGTCT	616
Qy	269	ProAlaGlyProTyrAlaLysGluSerGlyIleTrpAlaTyrTyrGluIleCys---Thr	287
Db	615	CCTCCTGGTTTATTACTGGTGAAGAAGTGTCTCTCATACATCGAATTTGTGTCTAGTTA	556
Qy	288	PheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValProTyrAlaTyr	307
Db	555	TTCCAGAAAAGAAATGGCATATTCAATACGATGAATATTACAAATGCTCCATACGGATAT	496
Qy	308	GlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTrp	327
Db	495	AATGATAAAATCTGGGTGGTGTACGAGATCTCGCTAGTATATCATGCAAGTTGGCCTTT	436
Qy	328	LeuLysHisAsnLysPheGlyAlaIleMetValTrpAlaIleAspLeuAspAspPheThr	347
Db	435	CTCAAGAATTTGGCGTCTCTGGCGTTATGATATGTCATTGGAACACGATGATTTCAA	376
Qy	348	GlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAlaLeuGly	367
Db	375	GGT---CATTCGGGACCG---AAATATCATCTTGTGAACAAAGTTCACATATGATCAAT	322



; TITLE OF INVENTION: AND USES THEREOF  
 ; FILE REFERENCE: AL-2-C3  
 ; CURRENT APPLICATION NUMBER: US/10/218,743  
 ; CURRENT FILING DATE: 2002-08-13  
 ; PRIOR APPLICATION NUMBER: US/09/292,225  
 ; PRIOR FILING DATE: 1999-04-15  
 ; PRIOR APPLICATION NUMBER: 60/098,909  
 ; PRIOR FILING DATE: 1998-09-02  
 ; PRIOR APPLICATION NUMBER: 60/085,295  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/098,565  
 ; PRIOR FILING DATE: 1998-04-17  
 ; PRIOR APPLICATION NUMBER: 09/062,013  
 ; PRIOR FILING DATE: 1998-04-17  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 39  
 ; LENGTH: 1527  
 ; TYPE: DNA  
 ; ORGANISM: Dermatophagoides farinae  
 US-10-218-743-39

Alignment Scores:  
 Pred. No.: 2,638-72 Length: 1527  
 Score: 711.50 Matches: 149  
 Percent Similarity: 54.17% Conservative: 72  
 Best Local Similarity: 36.52% Mismatches: 152  
 Query Match: 28.39% Indels: 35  
 DB: 9 Gaps: 11

US-10-004-219B-14 (1-455) x US-10-218-743-39 (1-1527)

QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21  
 DB 1428 CGAATCGTATGTATGTGGACATGGTCCGTTTATCATAA---GTTGATCCATACACA 1372  
 QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41  
 DB 1371 ATTGAAGATATTGATCCTTCAAAATGATCTCATTTGATGATGGTTTGTCTAAATTCGAT 1312  
 QY 42 AsnAsnGluIleThrThr-----IleGluTrp 50  
 DB 1311 GAATCAAAATACACCATTCATCAAGTTTGTGATCCATTCATCAAGATGATAACATCAATGG 1252  
 QY 51 AsnAspValThrLeuTyrGlnAlaPheAsnGlyLeuLysAsnLysSerGlnLeuLys 70  
 DB 1251 ---GAAACACCGGTATGACGTTTCAACACTTGAGATTGAGAAATCCAGAAATGACC 1195  
 QY 71 ThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrAlaProPheThrAlaMetValSer 90  
 DB 1194 ACCATGATTTCATTTGGTGGTATGAAGGTTTCAGAAAAATATTTCGGATATGGCAGCC 1135  
 QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlu 110  
 DB 1134 AATCCAAACATATCGTCAGCAATTTGTTCAATCAGTTTTCAGACTTTTTCAGAAATACAAA 1075  
 QY 111 PheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg---GlySerProProGln 129  
 DB 1074 TTCATGCTGCTAGATTGATTTGGGAAATATCTCGATCACTGTTAGGCAATCTTAATC 1015  
 QY 130 AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149  
 DB 1014 GATAAACAAACTATTAACTATTAGTAGAAGCTTAAAGAGGCAATTTGAACTTTTCGCGC 955  
 QY 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169  
 DB 954 -----TACTTGTGACTGCGCAGTATCACCCGGTAAAGATAAA 916  
 QY 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet 189  
 DB 915 ATTGACGTAGCTTATGAGCTCAAGAAATGAACCAATTTGTCATGTTGATGATGTCATG 856  
 QY 190 ThrTyrAspLeuHisGlySerTrpGluGlyTyrThrThrGlyLeuAsnSerProLeuTyrLys 209

DB 855 ACTTATGATTACCATGGCGGATGGGAAATGTTTTCGGCCATAATGCTCCGTTGTATATA 796  
 QY 210 TyrProThrAspThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyr 228  
 DB 795 CGACCCGATGAACGGATGAATTCACACTTACTTCAATGCACTACACCATGACCTAT 736  
 QY 229 TrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHis 248  
 DB 735 TATTTGAACAATGGCGCTACTCGAGACAACTTGTATGGGTGTTCCATTTCTATGGTGT 676  
 QY 249 AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268  
 DB 675 GCTTGGAGCATCGAAGATCGAAGCAAACTCAAACTTGGCGATCCGCCCAAAGGCATGCT 616  
 QY 269 ProAlaGlyProTyrAlaLysLysSerGlyIleTyrPalaTyrTyrGluIleCys---Thr 287  
 DB 615 CTCTCTGTTTTTATTACTGGTGAAGAGGTGTTCTCTCATACATCGAATTTGTGTCAGTTA 556  
 QY 288 PheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGlnGluValProTyrAlaTyr 307  
 DB 555 TTCCAGAAAGAAGATGCGCATATTCAATACGATGAATATTACAATGCTCCATACGGATAT 496  
 QY 308 GlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTrp 327  
 DB 495 AATGATAAAATCTGGTGTGGTACGATGATCTGGCTAGTATATCATGCAAGTTGGCCTTT 436  
 QY 328 LeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAspAspPheThr 347  
 DB 435 CTCAGAAATTTGGGCTCTCTGGGCTTATGATATGTTGATGGGCAAAACGATGATTTCAAA 376  
 QY 348 GlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAlaLeuGly 367  
 DB 375 GGT---CATTTGGGACCG---AATATCCATTTGTTGAACAAAGTTTCAATATGATCAAT 322  
 QY 368 -----LeuGlnSerAlaSerCysThrAlaProAlaGln 378  
 DB 321 GGTGATGAAAGAACTCTTAGCAATGCTTTTGGGCCCAAGTACAAACACCAACA--- 265  
 QY 379 ProIleGluProIleThrAlaAla 386  
 DB 264 CAACCCCGCTCAACTACTTCG 241

# RESULT 13

US-10-218-743-34  
 ; Sequence 34, Application US/10218743  
 ; Publication No. US20030096779A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McCall, Catherine A.  
 ; APPLICANT: Hunter, Shirley Wu  
 ; APPLICANT: Weber, Eric R.  
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS  
 ; TITLE OF INVENTION: AND USES THEREOF  
 ; FILE REFERENCE: AL-2-C3  
 ; CURRENT APPLICATION NUMBER: US/10/218,743  
 ; CURRENT FILING DATE: 2002-08-13  
 ; PRIOR APPLICATION NUMBER: US/09/292,225  
 ; PRIOR FILING DATE: 1999-04-15  
 ; PRIOR APPLICATION NUMBER: 60/098,909  
 ; PRIOR FILING DATE: 1998-09-02  
 ; PRIOR APPLICATION NUMBER: 60/085,295  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/098,565  
 ; PRIOR FILING DATE: 1998-04-17  
 ; PRIOR APPLICATION NUMBER: 09/062,013  
 ; PRIOR FILING DATE: 1998-04-17  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 34  
 ; LENGTH: 1621  
 ; TYPE: DNA  
 ; ORGANISM: Dermatophagoides farinae  
 ; FEATURE:





```
QY 42 AsnAsnGluIleThr-----IleGluTrp 50
Db 1392 GAATACAAATACACCACTCAAGTTTTCATCCATTTCAAGATGATAACCAATCACTG 1333
QY 51 AsnAspValThrLeuTyrglnAlaPheAsnGlyLeuLysAsnLysAsnSerGlnLeuLys 70
Db 1332 ---GAAACACCGGCTATGAAGCTTTCAACACTTTGAGATTGAAGATCCAGAAATGACC 1276
QY 71 ThrLeuLeuAlaIleGlyTrpAsnPheGlyThrAlaProPheThrAlaMetValSer 90
Db 1275 ACCATGATTTCATCGGCTGTTGGTATGAGGTTTCAAGAAATATTCGATATGGAGCC 1216
QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTrpGlu 110
Db 1215 AATCCAACATATCGTCAGCAATTTGTTCAATCAGTTTGGACTTTTGGCAAGATACAAA 1156
QY 111 PheAspGlyLeuAspPheAspTrpGluTyrglnTyrglnTyrglnTyrglnTyrgln 129
Db 1155 TTCGATGGCTAGATTGGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 1096
QY 130 AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149
Db 1095 GATAACAAACTATTAACTATTGATAGAGAACTTAAAGAGCAATTTGAACCTTTCGGC 1036
QY 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169
Db 1035 -----TACTTGTGTGACTGCGCAGTATCACCCGGTAAAGATAA 997
QY 170 IleGlnSerGlyTyrglnLeuProGlnLeuSerGlnTyrglnLeuAspTyrglnHisValMet 189
Db 996 ATTGACGTAGCTTATGAGCTTCAAGAAATGAACCAATTTGTCGATTGGATGAATGTCATG 937
QY 190 ThrTyrglnLeuHisGlySerTrpGluGlyTyrglnTyrglnTyrglnTyrglnTyrgln 209
Db 936 ACTTATGATTACCATGCGGATGGGAATGTTTCGGCCATAATGCTCCGCTGGTATATA 877
QY 210 TyrProThrAspThrGlySer---AsnAlaTyrglnLeuAsnValAspTyrglnValMetAsnTyrgln 228
Db 876 CGACCCGATGAACCGATGAATGTCACACTTACTTCAATGTCACTACACCATGCACTAT 817
QY 229 TrpLysAspAsnGlyAlaProAlaGlnLysLeuIleValGlyPheProThrTyrglnHis 248
Db 816 TATTTGAACAATGGGCTACTCTGAGCAAACTTGTATGGGTGTTCCATTCCTATGCTGT 757
QY 249 AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268
Db 756 GCTTGGAGCATCGAAGATCGAAGCAAGTCAAACTTGGCGATCCGCGCAAGGCAATGCT 697
QY 269 ProAlaGlyProTyrglnAlaLysGluSerGlyIleTrpAlaTyrglnTyrglnTyrgln 287
Db 696 CCTCTGCTGTTTATCTGCTGAGAGGTTCTCTCATACATCGAATGTGTGAGTTA 637
QY 288 PheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGlnGluValProTyrglnAlaTyrgln 307
Db 636 TTCCGAAAGAAGATGGCATATTCATACGATGAATATTAATGCTCTCATACGATAT 577
QY 308 GlnGlyAsnValTrpValGlyTyrglnAsnIleLysSerPheAspIleLysAlaGlnTrp 327
Db 576 AATGATAAATCTGGTGTGTTACCATGATCTGGGTAGTATATCATGCAAGTTGGCCTT 517
QY 328 LeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAspAspPheThr 347
Db 516 CTCAGAAGATTTGGCGCTCTCTGCGCTTATGATATGTCATTCATGAAACGATGTTCAA 457
QY 348 GlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAlaLeuGly 367
Db 456 GGT---CATTTGGCGCGCG---AAATATCATTTGTTGAACAAAGTTCACAAATATGATCAAT 403
QY 368 -----LeuGlnSerAlaSerCysThrAlaProAlaGln 378
Db 402 GGTGATGAAGAAGACTTTCAGAAATGCTTTTGGGCCCAAGTACACCAACCAACA--- 346
```

```
QY 379 ProIleGluProIleThrAlaAla 386
Db 345 CCAACCAACCGCTCACTACTCTG 322
RESULT 15
US-10-218-743-20
; Sequence 20, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1608)
US-10-218-743-20
Alignment Scores:
Pred. No.: 6,35e-72 Length: 1608
Score: 708.50 Matches: 149
Percent Similarity: 53.98% Conservative: 75
Best Local Similarity: 35.90% Mismatches: 154
Query Match: 28.27% Indels: 37
DB: 9 Gaps: 10
US-10-004-219B-14 (1-455) x US-10-218-743-20 (1-1608)
QY 2 GlnLeuThrCysTyrglnPheThrAsnTrpAlaGlnTyrglnProGlyLeuGlyArgPheMet 21
Db 43 AGAATTTGTTGTTATGTTGGAACATGTCCTGATATCATATAA---GTTGATCCATCACT 99
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrglnAlaPheAlaGlyArgGln 41
Db 100 ATCGAAGATATGATTCATTCATCAAGTGATACATTAATGATGTTTCGCTAAATTCAT 159
QY 42 AsnAsnGluIleThrThr-----IleGluTrp 50
Db 160 GAATACAAATACCAATCAAGTTTTCATCTTACCAAGATGATAACCAATCACTG 219
QY 51 AsnAspValThrLeuTyrglnAlaPheAsnGlyLeuLysAsnLysAsnSerGlnLeuLys 70
Db 220 ---GAAACACCGGCTTATGAAGCTTTCAACACTTTCGATTCGATTCGATTCGATTCG 276
QY 71 ThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrAlaProPheThrAlaMetValSer 90
Db 277 ACCATGATTTCATCGGCTGTTGGTATGAGGCTCGGAAATATTCGATATGCTGCTCA 336
QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTrpGlu 110
Db 337 AATCCACATATCGTCAACAATTCATCAATCAGTTTGGACTTTTTCGACAGATACAG 396
QY 111 PheAspGlyLeuAspPheAspTrpGluTyrglnTyrglnTyrglnTyrglnTyrgln 129
```

Db 397 TCGACGGTCTAGATTGGATTGGAGTATCTCGATTGGGTAAACCGAAATC 456  
Qy 130 AspLysHisLeuPheThrValLeuGlnGluMetArgGluAlaPheGluGlnGluAla 149  
Db 457 GATAAACAACAACTATTGGCTTTGGTTAGAGAACTTAAAGACGCTTTGAACCTCATGGC 516  
Qy 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169  
Db 517 -----TACTTGTTCAGTCTGCTGCAGTATCACCAGGTAAAGACAA 555  
Qy 170 IleGlnSerGlyTyrGluLeuProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet 189  
Db 556 ATCGACCGAGCTTATGATATCAAGAATGAACAAATTTGCGATTGGATGAATGTCTATG 615  
Qy 190 ThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209  
Db 616 ACATATGATTACCCAGGTGGATGGGAAACTTTTACGGTCACAAATGCTCCGTGTATATAA 675  
Qy 210 TyrProThrAspThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyr 228  
Db 676 CGACCAAGATCAAACTGATGATTGCACACTTACTTCAATGTCAACTACACCATGCACTAT 735  
Qy 229 TrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHis 248  
Db 736 TATTGCAACAATGGTGCACAGACAAATGGTAATGGGTGTTCCATTCCTATGGCCGT 795  
Qy 249 AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268  
Db 796 GCTTGGAGCATTTGAAGATCGAAGCAAACTCAAACTTTGGAGATCCAGCCAAAGGCATGCG 855  
Qy 269 ProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGluIleCys---Thr 287  
Db 856 CCCCCAGGTTTCACTTCTGGTGAAGAGGTGCTCTCATATATAGAAATGTGTCATTG 915  
Qy 288 PheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValProTyrAlaTyr 307  
Db 916 TTTCAAAAGAGAAGATGGCATATCCAATACGATGAATATTACAATGCTCCATATGGTTAC 975  
Qy 308 GlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTrp 327  
Db 976 AATGATAAAATCTGGTTCGGTTACGATGATCTGCCAGTATATCATGCAAGTTGGCTTTC 1035  
Qy 328 LeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAspPheThr 347  
Db 1036 CTGAAGAATTAGCGCTTCTCGTGTGTCATGTTTGTGTCATTGGAATAATGATGATTTCAA 1095  
Qy 348 GlyThrPheCysAsnGlnGlyLysPheProLeuIle----- 359  
Db 1096 GGT---CACTGCGGACCG---AAAAATCCATTGTTGAACAAAGTTCATATATGATTAAT 1149  
Qy 360 -----SerThrLeuLysLysAlaLeuGlnSerAlaSerCysThrAla 375  
Db 1150 GCGGATGAAGAAGAACTTTTCGAATGCAATTTGGTCCAGTACACGACCACTCCCA 1209  
Qy 376 ProAlaGlnProIleGluProIleThrAlaAlaProSerGlySer 390  
Db 1210 ACGACGACCCACACCCCGACTACACGCCCAACTCTTCT 1254

Search completed: June 30, 2003, 03:58:41  
Job time : 190.913 secs



Matches 201; Conservative 65; Mismatches 122; Indels 13; Gaps 7;  
QY 1 YOLTCYFTWNAQVRPGGRFMPDNIDPCLTLLIYAFAGRONNEITTIEMNDV-TLYQAP 59  
Db 22 YKLVCYFTWNAHSPGPASIMPHDLDFFLCTLLIFAFASNNQIVAKNLQENVDLYPEF 81  
QY 60 NGLKKNKNSOLKTLTLLAGWNGFPTAPMTAVSPENRQTITTSVIKFLROYEFDGLDFWE 119  
Db 82 NXLKERNRELKTLISIGWNGFGRSRTFMTLSTLANKEKIDSVISFLRHGFDGLDLPFL 141  
QY 120 YGSRGSPQDKHLFTVLVOEMREAFQBAKQINKPRLMVTAAVAGISNIQSGYEIPOL 179  
Db 142 YPGLRGSPHPRNFWLFLIEELQAFERBALLTQHPRLLLSAAVSGIPSIHTSDALL 201  
QY 180 SQYLDYIHWMTYDLHGSWEGYTGENSEPLKYPTDTCGSNAYLVNDVYVWYKONGAPAEKL 239  
Db 202 GRLDLPINVLSDYDLHGSWEKFTGHNSPLSLPESDKSSA-----YAMNWRKLGTPADKL 256  
QY 240 IVGFPTTYGHNFIILSNPNTGIGAPTSAGAPG--YAKESGIWAWYVEICTFLKNGATQGW 297  
Db 257 INGFTTYGHNFIILSKESKNGL--QTASMGSPGKYTKQAGFLAYVEVCSFVQR-AKGIW 313  
QY 298 DAPQEVPIYQGVWVYGVYVNIKSFIDIKAOWLKHNKFGAMWAIDLDDFTGTCNQKPP 357  
Db 314 IDYQYVYAFKGEWLGIDYDTISFSYKAMYVREHFGAMVWTLMDDDVRGTCGNGPFP 373  
QY 358 LI-STLKKALGLOSACTAPAQPIETIPITAPSGSGSGSS 398  
Db 374 LVHILNELL-VOTESNSTPL-POFWETSSVNASGPGSENTA 412

## RESULT 2

C31L2 HUMAN STANDARD; PRT: 383 AA.  
AC F36222; P30923;  
DT 01-JUN-1993 (Rel. 26, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-39) (39 kDa synovial protein) (YKL-40).  
GN CH13L1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Cartilage;  
RX MEDLINE=94064658; PubMed=8245017;  
RA Hakala B.E., White C., Recklies A.D.;  
RT "Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family.";  
RT J. Biol. Chem. 268:25803-25810(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Blood;  
RX MEDLINE=97386591; PubMed=9244440;  
RA Rehli M., Krause S.W., Andresen R.;  
RT "Molecular characterization of the gene for human cartilage gp-39 (CH13L1), a member of the chitinase protein family and marker for late stages of macrophage differentiation.";  
RL Genomics 43:221-225(1997).  
RN [3]  
RP SEQUENCE OF 22-45;  
RX MEDLINE=90328983; PubMed=2375755;  
RA Nyirkos P., Golds E.E.;  
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine mammary protein expressed during the non-lactating period.";  
RL Biochem. J. 269:265-268(1990).  
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.  
CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG, PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.  
CC -!- PTM: GLYCOSYLATED.  
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M80927; AAA16074.1; -  
DR EMBL; Y08374; CAA69661.1; -  
DR EMBL; Y08375; CAA69661.1; JOINED.  
DR EMBL; Y08376; CAA69661.1; JOINED.  
DR EMBL; Y08377; CAA69661.1; JOINED.  
DR EMBL; Y08378; CAA69661.1; JOINED.  
DR PIR; S10677; S10677.  
DR PIR; A33162; A33162.  
DR Genew; HGNC:1932; CH13L1.  
DR MIM; 601525; -  
DR InterPro; IPR001579; Chitinase 18/2.  
DR InterPro; IPR001223; Glyco\_hydro 18.  
DR Pfam; PF00704; Glyco\_hydro 18; 1.  
DR ProDom; PD000471; Glyco\_hydro 18; 1.  
DR PROSITE; PS01095; CHITINASE\_18; FALSE\_NEG.  
KW Glycoprotein; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 383  
FT CARBOHYD 60 60  
FT SEQUENCE 383 AA; 42613 MW; 76ADD8298EEC2D1 CRC64;  
Query Match 39.0%; Score 976.5; DB 1; Length 383;  
Best Local Similarity 51.2%; Pred. No. 4.9e-64;  
Matches 188; Conservative 62; Mismatches 108; Indels 9; Gaps 5;  
QY 1 YOLTCYFTWNAQVRPGGRFMPDNIDPCLTLLIYAFAGRONNEITTIEMNDVTLYQAPN 60  
Db 22 YKLVCYFTWNAHSPGPASIMPHDLDFFLCTLLIFAFASNNQIVAKNLQENVDLYPEF 81  
QY 61 GLKKNKNSOLKTLTLLAGWNGFPTAPMTAVSPENRQTITTSVIKFLROYEFDGLDFWE 120  
Db 82 TLKERNPNLKTLLSVGGWNGFGRSRTFMTLSTLANKEKIDSVISFLRHGFDGLDLPFL 141  
QY 121 PGSRGSPPQDKHLFTVLVOEMREAFQBAKQINKPRLMVTAAVAGISNIQSGYEIPOL 180  
Db 142 PGRS-----DKQHTTLLIKENKAEFIKEA-QPGKKQLLSAALSAGKVTIDSSYIAKIS 195  
QY 181 QYLDYIHWMTYDLHGSWEGYTGENSEPLKYPTDTCGSNAYLVNDVYVWYKONGAPAEKL 240  
Db 196 QHLDYISIMTYDFHGAWRGTTGHHSPPLRGQSDASPRFSNTDYAVGYMLRIGAPASKLV 255  
QY 241 VGFPTYGHNFIILSNPNTGIGAPTSAGAPGAGPAGYAKESGIWAWYVEICTFLKNGATQGW 300  
Db 256 MGIFPTGSRFTLAS-SETGVGAPISGPGIPGRFTKEAGTLAYEICDFEL-GATVHRTLG 313  
QY 301 QEVPIYQGVWVYGVYVNIKSFIDIKAOWLKHNKFGAMWAIDLDDFTGTCNQK- KFPIL 359  
Db 314 QQVPYATKGNQWVGYDDQESVKSVQYLLKDRQLAGAMVWALDDDFQGSFCQDLRFP 373  
QY 360 STLKKAL 366  
Db 374 NAIKDAL 380  
RESULT 3  
C31L2 HUMAN STANDARD; PRT: 390 AA.  
ID C31L2 HUMAN  
AC Q15782; Q15783; Q15749;  
DT 16-OCT-2001 (Rel. 40, Created)





Query Match 38.3%; Score 959.5; DB 1; Length 678;  
 Best Local Similarity 49.3%; Pred. No. 1.8e-62;  
 Matches 187; Conservative 56; Mismatches 129; Indels 7; Gaps 3;

QY 1 YLTCTYTNWAGYRPGCLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIWMD-VTLYQAF 59  
 DB 22 HKLVCTFTNWAHSRPGPASIILPHDLDPFLCTHLIFAFASMMNNQIVAKDLODEKILYPEF 81

QY 60 NGLKNKNSQLKTLAIGGNWFGTAPTMVSTPENROTFTITSVIFKLYOYFEGDGLDFWE 119  
 DB 82 NKLKERNRELKTLISIGGNWFGTSFTTMSLFANREKFIASVISLRLTHDFDGLDFPL 141

QY 120 YPGSRGSPQDKHLFTVLVQEMREAFEOBAKOINPRLMTAAVAGISNIQSGYEIPOL 179  
 DB 142 YPGLRGSPMHDRTWFLFLIEELLFAPRKEALLTMRPRLLSAAVSGVPHIVQTSYDVRL 201

QY 180 SQYLDYIHWYDYLHGSWEGYTGNSPLYKYPTDGTGSNAYLNVDYVMYWKDNGAPAEKL 239  
 DB 202 GRLLDFINVLSDYDLHGSWERFTGHNPLSPFDPKSSA-----YAMNYWRKLGAPSEKL 256

QY 240 IVGFPTYGHNFTILSNPSNTGICAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 299  
 DB 257 IMGIPYGTFTLLKASKNGLOARIGPASPGRKYTKQEGFLAYFELCSFW-GAKKHWD 315

QY 300 PQEVPIYQGNVWGYDNIKSFIDIKAOMLKHNKFGGAMWAIDLDDFTGTCNQKGFPLI 359  
 DB 316 YQYVPYANKGKWEVGYDIAISFYKAWFIRREHFGGAMVWTLDMDDVRGTGCTGPFPLV 375

QY 360 STLKALGLQASCTAPAQ 378  
 DB 376 YVINDILVRAEFSSTSLPQ 394

## RESULT 6

OGP\_PAPAN STANDARD; PRT; 623 AA.  
 AC P36718;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
 DE (Oviductin) (Estrogen-dependent oviduct protein).  
 GN OVGPI OR OGP.  
 OS Papio anubis (Olive baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Papio.  
 OX NCBI\_TaxID=9555;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oviduct;  
 RX MEDLINE=98244335; PubMed=9584944;  
 RA Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,  
 RA Donnelly K.M., Arias E.B., Jaffe R.C.;  
 RT "The baboon oviduct: Characteristics of an oestradiol-dependent  
 RT oviduct-specific glycoprotein."  
 RL Hum. Reprod. Update 3:541-552(1997).  
 RN [2]  
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
 RC TISSUE=Oviduct;  
 RX MEDLINE=91367180; PubMed=1716345;  
 RA Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,  
 RA Jaffe R.C.;  
 RT "Cloning of a recombinant complementary DNA to a baboon (Papio  
 RT anubis) estradiol-dependent oviduct-specific glycoprotein."  
 RL Mol. Endocrinol. 5:356-364(1991).  
 CC -!- FUNCTION: BINDS TO OCYTOE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
 CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.  
 CC -!- SUBCELLULAR LOCATION: Secretory granules.  
 CC -!- TISSUE SPECIFICITY: OVIDUCT.  
 CC -!- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M59903; AAB39765.1; -;  
 DR PIR; A37954; A37954;  
 DR InterPro; IPR001579; Chitinase\_18/2.  
 DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
 DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
 DR PROSITE; PS01095; CHITINASE\_18; FALSE\_NEG.  
 KW Glycoprotein; Fertilization; signal  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT CHAIN 22 623 OVIDUCT-SPECIFIC GLYCOPROTEIN.  
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 623 AA; 69291 MW; 9E21CE481FFF1268 CRC64;

Query Match 38.1%; Score 954.5; DB 1; Length 623;  
 Best Local Similarity 48.8%; Pred. No. 3.6e-62;  
 Matches 185; Conservative 59; Mismatches 128; Indels 7; Gaps 3;

QY 1 YLTCTYTNWAGYRPGCLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIWMD-VTLYQAF 59  
 DB 22 HKLVCTFTNWAHSRPGPASIILPHDLDPFLCTHLIFAFASMMNNQIVAKDLODEKILYPEF 81

QY 60 NGLKNKNSQLKTLAIGGNWFGTAPTMVSTPENROTFTITSVIFKLYOYFEGDGLDFWE 119  
 DB 82 NKLKERNRELKTLISIGGNWFGTSFTTMSLFANREKFIASVISLRLTHDFDGLDFPL 141

QY 120 YPGSRGSPQDKHLFTVLVQEMREAFEOBAKOINPRLMTAAVAGISNIQSGYEIPOL 179  
 DB 142 YPGLRGSPMHDRTWFLFLIEELLFAPRKEALLTMRPRLLSAAVSGVPHIVQTSYDVRL 201

QY 180 SQYLDYIHWYDYLHGSWEGYTGNSPLYKYPTDGTGSNAYLNVDYVMYWKDNGAPAEKL 239  
 DB 202 GRLLDFINVLSDYDLHGSWERFTGHNPLSPFDPKSSA-----YAMNYWRKLGAPSEKL 256

QY 240 IVGFPTYGHNFTILSNPSNTGICAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 299  
 DB 257 IMGIPYGTFTLLKASKNGLOARIGPASPGRKYTKQAGFLAYFELCSFW-GAKKHWD 315

QY 300 PQEVPIYQGNVWGYDNIKSFIDIKAOMLKHNKFGGAMWAIDLDDFTGTCNQKGFPLI 359  
 DB 316 YQYVPYANKGKWEVGYDIAISFYKAWFIRREHFGGAMVWTLDMDDVRGTGCTGPFPLV 375

QY 360 STLKALGLQASCTAPAQ 378  
 DB 376 YVINDILVRAEFSSTSLPQ 394

## RESULT 7

OGP\_BOVIN STANDARD; PRT; 537 AA.  
 AC Q28042;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
 DE (Oviductin) (Estrogen-dependent oviduct protein) (fragment).  
 GN OVGPI OR OGP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-47.

RC TISSUE=Oviduct; MEDLINE=94257768; PubMed=8199272;  
RX Sendai Y., Abe H., Kikuchi M., Satch T., Hoshi H.;  
RA "Purification and molecular cloning of bovine oviduct-specific  
RT glycoprotein";  
RL Biol. Reprod. 50:927-934 (1994).  
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.  
CC -!- SUBCELLULAR LOCATION: Secretory granules.  
CC -!- TISSUE SPECIFICITY: OVIDUCT.  
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D16639; BAA04065.1; -  
DR InterPro; IPR001579; Chitinase 18/2.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR PROSITE; PS01095; CHITINASE 18; FALSE\_NEG.  
KW Glycoprotein; Fertilization; Signal.  
FT NON TER 1  
FT SIGNAL <1 18  
FT CHAIN 19 537 OVIDUCT-SPECIFIC GLYCOPROTEIN.  
FT CARBOHYD 399 399 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 537 AA; 59617 MW; CFCDEE6F0212D791 CRC64;  
  
Query Match 37.6%; Score 943.5; DB 1; Length 537;  
Best Local Similarity 49.8%; Pred. No. 1.9e-61;  
Matches 182; Conservative 53; Mismatches 125; Indels 7; Gaps 3;  
  
QY 1 VOLTCTFTNWAQYRPGLRGFMNDPCLCTHLIYAFAGRQNNETITIEWND-VTLYQAF 59  
DB 19 HKLVCTFTNWAQYRPGASLPDLDFLCTHLVAFASNNQIVPKDPQDEKILYPEF 78  
  
QY 60 NGLKNNLSQKTLTAIGWNFGTAPTAMVSTENRQTITSVIKELRQVDFGLDFWE 119  
DB 79 NKLKERNGLKTLISLGGWNGFTVTRTTLSTFNSRERFVSSVIALRTHTGFDGLDFL 138  
  
QY 120 YGSRGSPQDKHLFTVLQEMREAEQAKQINKPLMTVAAGAAGISNIQSGYEIPOL 179  
DB 139 YPLGRGSPARDRTTFVLEELQAFKNEAQLTMRPRLLLSAAVSGDPHVQKAYEARLL 198  
  
QY 180 SOYLDYIHWMTYDLHGSWEGYTGENSEPLYKYPTDGTGSNAYLVNDVYVMYWKONGAPAEKL 239  
DB 199 GRLLDIFSLVSLDHLGSWEKVTGHNPSLPFGDPKSSA-----YANNYRQLGVPEKL 253  
  
QY 240 IVGFPTYGHNFILSNPNTGIGAPTSAGAPYAKESGIWAYYEICTFLKNGATGWDA 299  
DB 254 LMGLPYGRTFHLKASQNELRAQAQVSPGKYTKQAGFLAYYEICCFVRR-AKRWIN 312  
  
QY 300 PQEVPYAGNVWGVNDIKSPDIKAQWLKHNKFGGAMVWADLDPTCTFCNQGFPLI 359  
DB 313 DQVYPYAFKGEWGVYDDAISFGYKAFFIKREHFGGAMVWTLDDDFRGYFCGTGFPPLV 372  
  
QY 360 STLKAL 366  
DB 373 HTLNLL 179  
  
RESULT 8  
OGP SHEEP  
ID\_OGP\_SHEEP STANDARD; PRT; 539 AA.  
AC Q28542; Q28543;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
DE (Oviductin) (Estrogen-dependent oviduct protein) (Estrus-associated  
DE oviductal glycoprotein) (OEGP).  
GN OVGPI OR OGP.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-39.  
RX TISSUE=Oviduct; MEDLINE=95269691; PubMed=7750470;  
RA Desouza M.M., Murray M.K.;  
RT "An estrogen-dependent secretory protein, which shares identity with  
RT chitinases, is expressed in a temporally and regionally specific  
RT manner in the sheep oviduct at the time of fertilization and embryo  
RT development.";  
RL Endocrinology 136:2485-2496 (1995).  
RN [2]  
RP SEQUENCE OF 10-539 FROM N.A.  
RC STRAIN=Merino; TISSUE=Oviduct;  
RX MEDLINE=96329120; PubMed=8726871;  
RA Marshall J.T.A., Nancarrow C.D., Brownlee A.G.;  
RT "Cloning and sequencing of a cDNA encoding an ovine  
RT estrus-associated oviductal protein";  
RL Reprod. Fertil. Dev. 8:305-310 (1996).  
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.  
CC -!- SUBCELLULAR LOCATION: Secretory granules.  
CC -!- TISSUE SPECIFICITY: OVIDUCT.  
CC -!- DEVELOPMENTAL STAGE: LEVELS ARE HIGHEST IN THE FIMBRIA AND AMPULLA  
CC AT ESTRUS AND ON DAY 1 OF PREGNANCY, WHEN GAMETE TRANSPORT AND  
CC FERTILIZATION OCCURS IN THE E2-DOMINATED FALLOPIAN TUBE. LEVELS  
CC DECLINE SIGNIFICANTLY ON DAY 2 AND UNDERGO A FURTHER SIGNIFICANT  
CC REDUCTION ON DAY 3 OF PREGNANCY COINCIDENT WITH TRANSPORT OF THE  
CC EMBRYO FROM THE OVIDUCT TO THE UTERUS, A REPRODUCTIVE STAGE  
CC ASSOCIATED WITH RISING PROGESTERONE LEVELS.  
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U16719; AAC48471.1; -  
DR EMBL; U17988; AAB01052.1; -  
DR InterPro; IPR001579; Chitinase 18/2.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR PROSITE; PS01095; CHITINASE 18; FALSE\_NEG.  
KW Glycoprotein; Fertilization; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 539 OVIDUCT-SPECIFIC GLYCOPROTEIN.  
FT CARBOHYD 402 402 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 13 13 M -> V (IN REF. 2).  
FT CONFLICT 122 122 K -> N (IN REF. 2).  
FT CONFLICT 282 282 A -> V (IN REF. 2).  
FT CONFLICT 375 375 A -> V (IN REF. 2).  
FT CONFLICT 484 484 R -> H (IN REF. 2).  
FT CONFLICT 520 520 I -> T (IN REF. 2).  
SQ SEQUENCE 539 AA; 59535 MW; F35000269987C193 CRC64;  
  
Query Match 37.6%; Score 941.5; DB 1; Length 539;  
Best Local Similarity 48.8%; Pred. No. 2.7e-61;  
Matches 179; Conservative 59; Mismatches 122; Indels 7; Gaps 3;  
  
QY 1 VOLTCTFTNWAQYRPGLRGFMNDPCLCTHLIYAFAGRQNNETITIEWND-VTLYQAF 59  
DB 19 HKLVCTFTNWAQYRPGASLPDLDFLCTHLVAFASNNQIVPKDPQDEKILYPEF 78



Db 22 HKLVCFYTNWAFSRPGSASILPRDLDPFLCTHLVFAFASMMNNQIVPKDPLKILYPEF 81  
Qy 60 NGLKNKNSQLTKLLAIGGNWFGTAPFTAMVSTPENROTEITSVIRKLOYEDGLDFDWE 119  
Db 82 NKLKERNRGLTKLLSVGGNWFSTRTKMLSTFSNRRERVKSVIALLRHTRHGFGLDLPL 141  
Qy 120 YPGSRGSPQDKHLFTVLVQEMREAFEOBAKQINPKRLMVTAAVAAGISNIQSGYEIPOL 179  
Db 142 YPGLRGSPARDWTFVFLLEELLQAFKNEAQLTMRPRLLSAAVSGDPHVIQKAYDARLL 201  
Qy 180 SOYLDYIHWMTYDLHGSWEGYTGENSEPLYKYPTDTGNSNAYLNVDYVMYWKDNGAPAEKL 239  
Db 202 GRLDFISVLSYDLHGSWEKVTGHSNPLSFSLPGLDPKSSA-----YAMSVMROLGVPPEKL 256  
Qy 240 IVGFPTYGHNFILSNPSNTGIGAPTSAGAPGAPYAKESGIWAYYEICTFLKNGATQGWDA 299  
Db 257 LMGPTYGRTFHLRLASQNELGAGAAGPASPAGKYKQAGFLAYEVCVSFVOR-AKKRWIN 315  
Qy 300 POEVPYAYGNNVWGYDNTKSFIDIKAOHLKHNKFGGAMWAIDLDDFTTCNQGKFFPLI 359  
Db 316 DQYVPYAFKGEWVGYYDAISFGYKAFKIKRHFHGGAMVWTLDDDFRGNFGCTGPFPLA 375  
Qy 360 STLKAL 366  
Db 376 HTLNLL 382

RESULT 9  
OGP\_PIG STANDARD; PRT; 527 AA.  
AC Q28990;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)  
DE (Oviductin) (Estrogen-dependent oviduct protein) (POSP-E3).  
GN OVGPI OR OGP.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=Yorkshire X Duroc X Hampshire; TISSUE=Oviduct;  
RA MEDLINE=97107140; PubMed=894988;  
RA Buhi W.C., Alvarez I.M., Choi I., Cleaver B.D., Simmen F.A.;  
RT "Molecular cloning and characterization of an estrogen-dependent  
porcine oviductal secretory glycoprotein."  
RL Biol. Reprod. 55:1305-1314(1996).  
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE  
IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.  
CC -!- SUBCELLULAR LOCATION: Secretory granules.  
CC -!- TISSUE SPECIFICITY: OVIDUCT.  
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U43490; AAA85445.1;  
DR InterPro; IPR001579; Chitinase\_18/2.  
DR InterPro; IPR001223; Glyco\_hydro\_18.  
DR Pfam; PF00704; Glyco\_hydro\_18; 1.  
DR ProDom; PD000471; Glyco\_hydro\_18; 1.  
DR PROSITE; PS01095; CHITINASE\_18; FALSE\_NEG.  
KW Glycoprotein; Fertilization; Signal  
FT SIGNAL 1 21 BY SIMILARITY.  
FT CHAIN 22 527 OVIDUCT-SPECIFIC GLYCOPROTEIN.  
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 527 AA; 58519 MW; 3187849CA2363A2 CRC64;  
Query Match 37.4%; Score 936.5; DB 1; Length 527;  
Best Local Similarity 47.3%; Pred. No. 66-61;  
Matches 186; Conservative 61; Mismatches 123; Indels 23; Gaps 5;  
Qy 1 YOLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIEMNDVTL-YQAF 59  
Db 22 HKLVCFYTNWAFSRPGSASILPRDLDPFLCTHLVFAFASMMDSQIVAKDARDSEIFYPEF 81  
Qy 60 NGLKNKNSQLTKLLAIGGNWFGTAPFTAMVSTPENROTEITSVIRKLOYEDGLDFDWE 119  
Db 82 NKLKERNRGLTKLLSVGGNWFSTRTKMLSTFSNRRERVKSVIALLRHTRHGFGLDLPL 141  
Qy 120 YPGSRGSPQDKHLFTVLVQEMREAFEOBAKQINPKRLMVTAAVAAGISNIQSGYEIPOL 179  
Db 142 YPGLRGSPARDWTFVFLLEELLQAFKNEAQLTMRPRLLSAAVSGDPHVIQKAYDARLL 201  
Qy 180 SOYLDYIHWMTYDLHGSWEGYTGENSEPLYKYPTDTGNSNAYLNVDYVMYWKDNGAPAEKL 239  
Db 202 GRLDFISVLSYDLHGSWEKVTGHSNPLSFSLPGLDPKSSA-----YTMVYWRKLGAPPEKL 256  
Qy 240 IVGFPTYGHNFILSNPSNTGIGAPTSAGAPGAPYAKESGIWAYYEICTFLKNGATQGWDA 299  
Db 257 LMGPTYGRTFHLRLASQNELGAGAAGPASPAGKYKQAGFLAYEVCVSFVOR-AKKRWID 315  
Qy 300 POEVPYAYGNNVWGYDNTKSFIDIKAOHLKHNKFGGAMWAIDLDDFTTCNQGKFFPLI 359  
Db 316 DQYVPYAFKGEWVGYYDAISFGYKAFKIKRHFHGGAMVWTLDDDFRGNFGCTGPFPLV 375  
Qy 360 STLKAL 376  
Db 376 YMLNLLKAEVSSSTLSPGFLSTTVNSRTCP 408

RESULT 10  
CHIT CAEEL  
ID CHIT CAEEL STANDARD; PRT; 617 AA.  
AC Q11174;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Putative endochitinase (EC 3.2.1.14).  
GN C04F6.3  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Nhan M.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
acetyl-D-glucosamine polymers of chitin.  
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
HYDROLASES).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U42835; AAA83586.1;  
DR WormPep; C04F6.3; C030923.  
DR InterPro; IPR002557; Chitin\_bind\_Pera.  
DR InterPro; IPR001579; Chitinase\_18/2.  
DR InterPro; IPR001223; Glyco\_hydro\_18.

```
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR Pfam: PF01607; CBM 14; 2.
DR ProDom: PD000471; Glyco_hydro_18; 1.
DR SMART: SM00494; ChitB2; 2.
DR PROSITE: PS01095; CHITINASE_18; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Chitin degradation.
FT ACT_SITE 179 179 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 435 478 THR-RICH.
SQ SEQUENCE 617 AA; 66857 MW; DDA1D2AAACOE54DA CRC64;

Query Match 35.3%; Score 885; DB 1; Length 617;
Best Local Similarity 39.2%; Pred. No. 4.2e-57;
Matches 190; Conservative 80; Mismatches 169; Indels 46; Gaps 13;

QY 1 YQLTCVFTNWAQYRPGLRGFMFDNDPCLCTHLIYAFAGRQNNIIT-----IEKN 51
DB 53 YIRPCYFTNWAQYRQGRKAFVPEYDTPGCLTHLFAF-GWMAADYTVRAYDPAIDPNDWA 111
QY 52 DVTLYQAFNGLKNNKSQLTLLAIGWNFGTAPFTAMVSTPENRQFTITSVIFLQRYEF 111
DB 112 GEGMYRRVVKLKVTDQLATLLSFGCWSFGTALFQGMASASRKFVIDSAITFVRTWGF 171
QY 112 DGLDFWEYFGSRGSPQDKHLFTVLVQEMREAFEOAKQINPRLMVTAAVAAGISNIQ 171
DB 172 DGDIDWIEYF-----SGATDMANYVALVKELKAACESEAGSTGKDRLLVTAAVAAGPATID 227
QY 172 SGVEIPOLSOYLDYIHVMTYDLHGSWEGVTGENSPLY---KYPTDTGSNAYLNVDMY 228
DB 228 AGYDIENLAPNDFILMYSYDFPGAWASLVGNFSPLYATTELPAEWNG---WNVDSSARY 284
QY 229 WKDNGAPAEKLIYGPFTYGHNFILSPNFTGIGAPTSAGAPG--GPYAKESGIWAYYEIC 286
DB 285 WNKQGMPEKILVMPYTYGRGWTNNAS--AINPGTSGS-PAKITQYQVQEGVAGYFEC 341
QY 287 TFLKNGATCGWADAPQVPVAYQGNVWVDNITSKFDIKQWLKHNKFGGAMVWADLDDF 346
DB 342 EMLANGATRYWDSQSQVPLVQGNOWMSYDDDESFANKMAYVKREGYGFVMTLDFDDF 401
QY 347 TGTFCNOG--KPLISTLAKALG-----LQSASCTAP-----AQPTEPITAPSGS 390
DB 402 NAGCSNSNQLPLISVIAKELGGVILPKKGGVTTPATTTVTTGRP--PWTSATVTT 459
QY 391 GNGSGSSSSGG---SSGSGGFCVAVRNGLYPVANNRNAPFHWCVNGTYQONCOAGLVFDT 447
DB 460 TAATTTTTRAAATTTTASNTNVCSKSDGFVPNSNCGFLVFLCLSSKSYMSWCPGLOYS 519
QY 448 SCDC 452
DB 520 SLKYC 524

RESULT 11
C3L1_MOUSE
ID C3L1_MOUSE STANDARD; PRT; 381 AA.
AC Q61362;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE (GP-39) (BRP39 protein).
GN CH13L1 OR BRP39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=EVB/N; TISSUE=Breast;
RX MEDLINE=95060797; PubMed=7970700;
RA Morrison B.W., Leder P.;
RT "neu and ras initiate murine mammary tumors that share genetic markers
RT generally absent in c-myc and int-2-initiated tumors.";
RL Oncogene 9:3417-3426(1994).
```



```
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE 18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Antigen;
KW Repeat; Glycoprotein; Calcium-binding.
FT SIGNAL 1 22 ENDOCHITINASE.
FT CHAIN 23 504
FT DOMAIN 23 400 CATALYTIC.
FT DOMAIN 401 450 SER/THR-RICH (LINKER).
FT DOMAIN 407 448 3 X 14 AA APPROXIMATE TANDEM REPEATS.
FT ACT SITE 148 148 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 504 AA; 55971 MW; A78BE7BF8E3709B CR664;

Query Match 31.4%; Score 786.5; DB 1; Length 504;
Best Local Similarity 35.0%; Pred. No. 4.9e-50;
Matches 169; Conservative 73; Mismatches 200; Indels 41; Gaps 9;

QY 1 YQLTCYFTWAQYRPGCLGRPMEDNDPCLTCLHLYAFAG-RQNNETITIEWDVT-----54
DB 23 YRGCYTTWAQYRPGEGKFLPNTGNGICTHLYAFKVDLGDGSKPFEWDEDETSK 82
QY 55 -LYQAFNGLNKNSQLKTLAIGWNFGTAPFTAMVSTPENRQTFTTSVIKFLROYEPDG 113
DB 83 GMSYATKLRNTPGLKVLGSGYNGFSAITGTAKSAQKTERFIKSAIAFLRKNFPG 142
QY 114 LDFDWEYPSGRSPPODKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAGISNIQSG 173
DB 143 FLDLWEYVPG---VAEEH--AKLVEAMKTAFAVEAKTSGKQRLLLTAASVAGKGTIDGS 196
QY 174 YEIPQLSQVLDYHUMTYDLHGSWEYTGENSEPLYKYPTDTSNAYLNVDYVMYWKONG 233
DB 197 YNVESLGRNFDLLFLMSYDLHGSWEKNVDLHGKHPKTEGVSGIGIFNTEFAADYWASKG 256
QY 234 APAELIVGFPYVGHNFILSNPNTGIGAPTSGAGPAGYAKESGIWAYYEICTFLKNGA 293
DB 257 MPEKIIIGIPYACQWTLNDSETAIGAASRPSASAKTNPAGGTASTWEICKYLKEGG 316
QY 294 TQGDWAPQBYVAYQGNVWGVYDNKISFDIKAQWLKHNKFGAMWAIDLDFGTGFCNQ 353
DB 317 KETVHOEGVGAYVWKGQWGYDNEETIRIKMKWLKEGYGGAFTWALDFDFTGKSCGK 376
QY 354 GKFLISTLKKALGLOSASCTAPAOPIEP-ITAAPSGSGNGSGSS-----398
DB 377 GPYPLNLAISSELEGESENPEITTE--EPSITETAYETDETEETSETBAYDTDETEETS 434
QY 399 -----SGSGSGSGGFCVANGRIYVANNRNFAFWHCVNGVYVQCNQAGLGFDTSCD 450
DB 435 ETEATYTDTEEGQE--CPEP-DGLFPHTDCHLFIQCANIAYVMQCPATTFENDAIK 491
QY 451 CCN 453
DB 492 VCD 494

RESULT 14
CHIL_APHAL
ID CHIL_APHAL STANDARD; PRT; 423 AA.
AC P32470;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Chitinase 1 precursor (EC 3.2.1.14).
GN CHIL.
OS Aphanocladium album.
OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Aphanocladium.
OX NCBI_TaxID=12942;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ETHM 483;
RA MEDLINE=93013040; PubMed=1398137;
RX Blaiseau P.-L., Lafay J.-F.;
RT "Primary structure of a chitinase-encoding gene (chil) from the filamentous fungus Aphanocladium album: similarity to bacterial chitinases.";
```

CH11\_BACCI STANDARD; PRT; 699 AA.  
AC P20533;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Chitinase A1 precursor (EC 3.2.1.14).  
GN CHIA1.  
OS Bacillus circulans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1397;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WL-12;  
RX MEDLINE=90368776; PubMed=2203782;  
RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;  
RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed  
RT its evolutionary relationship to Serratia chitinase and to the type  
RT III homology units of fibronectin.";  
RL J. Biol. Chem. 265:15659-15665(1990).  
RN [2]  
RN MUTAGENESIS.  
RC STRAIN=WL-12;  
RX MEDLINE=93366760; PubMed=8103047;  
RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,  
RA Uchida M., Tanaka H.;  
RT "Identification of glutamic acid 204 and aspartic acid 200 in  
RT chitinase A1 of Bacillus circulans WL-12 as essential residues for  
RT chitinase activity.";  
RL J. Biol. Chem. 268:18567-18572(1993).  
CC -! CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
CC acetyl-D-glucosamine polymers of chitin.  
CC -! SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC -! SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
CC HYDROLASES).  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M57601; AAA81528.1; -;  
CC PIR; A38368; A38368.  
CC HSP; P02751; LFNA.  
CC InterPro; IPR003610; CBM\_5\_12.  
CC InterPro; IPR001579; Chitinase\_18/2.  
CC InterPro; IPR003961; FN\_III.  
CC InterPro; IPR003962; FNIII repeat.  
CC InterPro; IPR001223; Glyco\_hydro\_18.  
CC Pfam; PF00041; fn3; 2.  
CC Pfam; PF00704; Glyco\_hydro\_18; 1.  
CC Pfam; PF02839; CBM\_5\_12; 1.  
CC PRINTS; PRO0014; FNTYPEIII.  
CC ProDom; PD000471; Glyco\_hydro\_18; 2.  
CC SMART; SM00495; ChtBD3; 1.  
CC SMART; SM00060; FN3; 2.  
CC PROSITE; PS01095; CHITINASE\_18; 1.  
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Repeat.  
FT SIGNAL 1 41  
FT CHAIN 42 699  
FT DOMAIN 42 460  
FT CATALYTIC.  
FT DOMAIN 465 549  
FT FIBRONECTIN TYPE-III (R-1).  
FT FIBRONECTIN TYPE-III (R-2).  
FT PROTON DONOR (PROBABLE).  
FT ACET SITE 204 204  
FT D->N: DECREASE IN ACTIVITY.  
FT MUTAGEN 200 200  
FT D->E: NO CHANGE IN ACTIVITY.  
FT MUTAGEN 200 200  
FT E->D:Q: LOSS OF ACTIVITY.  
FT SEQUENCE 699 AA; 73677 MW; AC7C9B22E2987643 CRC64;  
Query Match 18.5%; Score 463.5; DB 1; Length 699;  
Best Local Similarity 27.9%; Pred. No. 2.6e-26;

Matches 129; Conservative 76; Mismatches 172; Indels 85; Gaps 17;  
QY 1 YQLTCYFTNWAQYRPLGR-FMPDNIDPCLCTHLIYAFAG----- 39  
DB 44 YKIVGYFSWAAY-----GRNVNADIDPTKTHINAFADICWNGIHGPNPDSGPNPTW 99  
QY 40 -RQNEITITIE-----WNDVTLYQA-----FNLKKNKNSQLKTL 73  
DB 100 TCQNEKSTINVPNGTIVLGDPIWDTKTFTAGDTWDQPIAGNINQLNKTKQTPNPKTII 159  
QY 74 AIGWNFTGAPTAMVSTPENRQTFITSVIKFLROYEEDGLDFDWEYPSRG-----SPP 128  
DB 160 SVGGWTSNR-FSDVAATAATREVFANSAVDPLRKYNFDGVDLDWEYPSVGGDGNKRP 218  
QY 129 ODKHLFTVLVOEMREAFQEAQINKPRLMTVAAGAISNIQSQYEIPQLSOYLDYIHV 188  
DB 219 EDKQNYTILLKIREKLDAGAVDCKYLLIYASGS--ATYAANTELAKIAIIVDWINI 276  
QY 189 MTYDLHGSWEGYTGENSEPLYKYPTDTGSNA-----YLNVDYVNMVYKDNAGAPAEKLIIVGPP 244  
DB 277 MTYDPNGAWQKISAHNAPLNVDPAASAAGVPDANTFNVAAGAQGHLDAGVPAKLVLP 336  
QY 245 TYGHNFILSNFTGIGAPTSAGAGPAGYAKESGIWAYYEI-CTFL-KNGATQGHDAPOE 302  
DB 337 FYGRGWDCAQAGNGQYQTCGSSVGTW--EAGSFDFYDLEANYINKNGYTRYWNTAK 394  
QY 303 VPYAYQGN--VMVGVDNFKSFDIKAOWLKHNKFGGAMWAIDDDFTGTFCNOGKFFPLIS 360  
DB 395 VPYLYNASKRFISYDDAESGYKTAYIKSKGLGGMFWELS-GDRNKTQLQKLADEL-- 451  
QY 361 TLKKALGLQASCTAPAPQPIBPITAAPSGSGNGSGSSSSSGS 402  
DB 452 -----PTGGTVP--PVD--TTAPSVPGNARSTGVTANS 480

Search completed: June 29, 2003, 21:00:23  
Job time : 14.0769 secs

**THIS PAGE BLANK (USPTO)**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 29, 2003, 21:10:28 ; Search time 1499.62 Seconds  
(without alignments)  
4913.891 Million cell updates/sec

Title: US-10-004-219b-14

Perfect score: 2506

Sequence: 1 YQLTCYFTNQAQYRPLGRF.....QQNCOAGLVFTSCDCNWA 455

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlp

-Q=/cgn2\_1/USPTO\_spool/US10004219/runat\_24062003\_160228\_8173/app.query.fasta\_1.1877  
-DB=EST -QFMT=FASTAP -SUFFIX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10004219 @CGN 1 1.4575 @runat\_24062003\_160228\_8173 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hcc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hcc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_nam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rod.\*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2097.5	83.7	1533	11	AK008633	AK008633 Mus muscu
2	2079.5	83.0	1535	11	AK008757	AK008757 Mus muscu
3	2035.5	80.8	1534	11	AK008650	AK008650 Mus muscu
4	1980	79.0	1530	11	AK007573	AK007573 Mus muscu
5	1438.5	57.4	957	13	BI760250	BI760250 603045190
6	1375	54.9	1105	14	BM923480	BM923480 AGENCOURT
7	1327.5	53.0	892	14	BQ231098	BQ231098 AGENCOURT
8	1270	50.7	838	12	BG867815	BG867815 602786336
9	1251	49.9	1038	13	BI517797	BI517797 603042049
10	1241	49.5	872	14	BQ231064	BQ231064 AGENCOURT
11	1237	49.4	809	12	BG866417	BG866417 602785352
12	1209	48.2	764	12	BG869304	BG869304 602788819
13	1207	48.2	763	12	BG872647	BG872647 602793705
14	1205	48.1	837	12	BG868741	BG868741 602787570
15	1202	48.0	939	14	BQ219632	BQ219632 AGENCOURT
16	1201	47.9	857	12	BG871434	BG871434 602790636
17	1198	47.8	894	12	BG869958	BG869958 602789826
18	1195.5	47.7	843	12	BG868207	BG868207 602784727
19	1194	47.6	1100	12	BF302298	BF302298 602031269
20	1188.5	47.4	855	12	BG872136	BG872136 602792974
21	1182	47.2	833	12	BG871337	BG871337 602790509
22	1181.5	47.1	949	12	BG868019	BG868019 602788381
23	1180.5	47.1	886	12	BG870840	BG870840 602791891
24	1178	47.0	824	13	BI518221	BI518221 603042049
25	1178	47.0	880	12	BG871468	BG871468 602790678
26	1172	46.8	847	12	BG866775	BG866775 602786605
27	1170.5	46.7	848	12	BG869143	BG869143 602789020
28	1170	46.6	856	12	BG870147	BG870147 602793390
29	1169	46.6	1155	12	BG867473	BG867473 602788576
30	1167.5	46.6	836	12	BG869655	BG869655 602789716
31	1167.5	46.6	976	14	BQ225554	BQ225554 AGENCOURT
32	1167	46.6	800	12	BG866853	BG866853 602786689
33	1164.5	46.5	788	12	BG873335	BG873335 602794304
34	1162.5	46.4	900	12	BG867646	BG867646 602787435
35	1161	46.3	830	12	BG869896	BG869896 602789435
36	1160	46.3	923	13	BI686954	BI686954 603313831
37	1159.5	46.3	882	12	BG873488	BG873488 602791730
38	1159	46.2	742	12	BG873143	BG873143 602794547
39	1157.5	46.2	863	12	BG872850	BG872850 602793373
40	1156.5	46.1	874	12	BG865756	BG865756 602784345
41	1156.5	46.1	948	13	BI664452	BI664452 603290028
42	1155	46.1	828	12	BG869811	BG869811 602789327
43	1151.5	45.9	845	12	BG865848	BG865848 602788264
44	1150	45.9	757	12	BG867292	BG867292 602788635
45	1150	45.9	837	12	BG865578	BG865578 602783726

ALIGNMENTS

RESULT 1  
AK008633  
LOCUS  
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:chitinase, acidic, full insert sequence.  
ACCESSION AK008633  
VERSION AK008633.1 GI:12842941  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain: C57BL/6J) adult male stomach cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library  
clone:2200003E03.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1





Db 260 GACTTGAATAAAGCAGGACAGCAAACTGAAACCCCTCTGGCAATTTGGAGGCTGGAACTTT 319  
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAenArgGlnThrPheIleThr 100  
Db 320 GGAAGCTGCTCTTTTCACTACCATGGTTTCCACTTCTCAGAACCCGACAGACCTTCATTACC 379  
Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
Db 380 TCAGTCATCAAAATTTCTGGTCAGTATGGTTTGTGACCTGGACCTGGAGCTGGGAATAC 439  
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
Db 440 CCAGGCTCAGCTGGAGCCCTCTCAGGACAGCATCTCTTCACTGTCTCTGTGGTGAAGAA 499  
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAenLysProArgLeuMetValThr 160  
Db 500 ATGCGTGAAGCTTTTGGAGGAGGCTATTGAGAGCAACAGGCCCGACGATGATGTTACT 559  
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
Db 560 GCTGCTGTAGCTGGTGGATTTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACCTTCT 619  
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
Db 620 AAGTACCTGGATTTTCACTCATGATCATATGACCTCCATGCTCTCTGGGAGGCTAC 679  
Qy 201 ThrGlyGluAenSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
Db 680 ACTGGGGAATAGTCCCTTTTACAAATACCTCTAGAGCTGGTGAAGATGCCCTACCTC 739  
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240  
Db 740 AATGTGGATTATGATGACTATTGGAGAACATGAGCCCGACCTGAGAGCTCAT 799  
Qy 241 ValGlyPheProThrTyrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyIle 260  
Db 800 GTTGGATTCCAGAGATATGGACACACCTTCATCTCGAGAAACCCCTCTCATTAATGCAATT 859  
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280  
Db 860 GTGTCCTTACCTCTCTGGTATGGCCCTCTGAGGCTTATACAGAGCTGGGTTCTGG 919  
Qy 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaPro 300  
Db 920 GCCTACTATGATTTGACCTTTCTGAGAGTGGAGCCACTGAGGTCTGGGATGCTCTCC 979  
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAenValTrpValGlyTyrAspAsnIleLysSer 320  
Db 980 CAAGAAGTGGCCCTATGCTTATAAGCCCAACGAGTGGCTTGGCTATGACAAATATCAAGAGC 1039  
Qy 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340  
Db 1040 TTCAGTGTAAAGGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGGGTGGCCATGATGGGCC 1099  
Qy 341 IleAspLeuAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360  
Db 1100 ATTGACCTTGATGACTTCACTGGCTTCTCTGTGATCAGGGAATTTCTCTGACTTCT 1159  
Qy 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380  
Db 1160 ACTTTGAAACAAAGCCCTTGGCATATCCACTGAAGGTTGACACAGCTCCTGACGTGCTTCC 1219  
Qy 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400  
Db 1220 GAGCCAGTGACTTCTCTCA-----GGAAGTGGGAGTGGGGGTGGAGCTCCGGA 1270  
Qy 401 GlySerSerGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420  
Db 1271 GGAAGCTCTGGAGGAGTGGATTTCTGTGCCGCAAGACGATGGCTCTTACCTCTGTGGCA 1330  
Qy 421 AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440  
.....

Db 1331 GATGACAGAAATGCTTTTGGCAGTCATCAATGGAATCATCATACCAGCAGCATTTGTCAA 1390  
Qy 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454  
Db 1391 GCAGGGCTTGTTTTGTATACCATGCTGTAAATGCTGCAACTGG 1432  
RESULT 2  
AK008757 1535 bp mRNA linear HTC 19-JAN-2002  
LOCUS Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210020N23.chitinase, acidic, full insert sequence.  
DEFINITION AK008757 GI:12843148  
VERSION AK008757.1  
KEYWORDS Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library  
SOURCE clone:2210020N23.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kaaukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Glissi, C., King, B., Koichiwa, H., Kuhl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., But, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Yoshinobori, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851  
5 (bases 1 to 1535)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,





## Alignment Scores:

Pred. No.: 1.32e-204 Length: 1534  
 Score: 2025.50 Matches: 366  
 Percent Similarity: 89.43% Conservative: 40  
 Best Local Similarity: 80.62% Mismatches: 45  
 Query Match: 80.83% Indels: 5  
 DB: 11 Gaps: 1

US-10-004-219B-14 (1-455) x AK008650 (1-1534)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
 DB 80 TACAATCTCATATGCTATTTTCAACCACTGGGCCAGTATCGGCCAGGCTGGGGAGCTTC 139  
 QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40  
 DB 140 AAGCCTGATGACATTAACCCCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199  
 QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60  
 DB 200 CAGAACATGAGATCACCAACCATAGATGATGATGATGATGATGATGATGATGATGATG 259  
 QY 61 GlyLeuLeuAsnIleAsnSerGlnLeuLeuThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
 DB 260 GACTTTGAAAAACAGGAACAGCAACCTCTCGCAATGGAAGCTGGAGCTTC 319  
 QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
 DB 320 GGAACCTGCTCTTCACTACCATGCTTCCACTTCTCAGAACCGCCAGACCTTCATTACC 379  
 QY 101 SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
 DB 380 TCAGTCATCAAAATTTCTGCTGATGATGCTTGTGATGATGATGATGATGATGATGATG 439  
 QY 121 ProGlySerArgGlySerProProGlnAsnAspLeuHisLeuPheThrValLeuValGlnGlu 140  
 DB 440 CCAGGCTCAGTGGGAGCCCTCTCTCAGGACCAAGCATCTCTTCACTGCTCTGGTGGAG 499  
 QY 141 MetArgGluAlaPheGluGlnGluAlaLeuGlnIleAsnLeuProArgLeuMetValThr 160  
 DB 500 ATGCTGGAAGCTTTTGAAGGAGGCTATTGAGACCAAGCCCAAGCATGATGATGATG 559  
 QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
 DB 560 GCTGCTGTAGTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACCTTCT 619  
 QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
 DB 620 AAGTACCTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATG 679  
 QY 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
 DB 680 ACTGGGGA-AATAGTCTCTTTACAAATACCTCTAGACTGGTATGATGATGATGATGATG 738  
 QY 221 AsnValAspTyrValMetAsnTyrTrpIleAspAsnGlyAlaProAlaGlyLeuIle 240  
 DB 739 AATGTGATATGTCATGAATATTGGAAGAACATGAGGCCCAAGCTCAGAAAGCTCAT 798  
 QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260  
 DB 799 GTTGGATC-CGAGAGATGATGACACACTTTCATCTCTGAGAAACCCCTCTGATATGGA 857  
 QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLeuGlySerGlyIleTrp 280  
 DB 858 GTTGGCCCTACTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917  
 QY 281 AlaTyrTyrGluIleCysThrPheLeuLeuAsnGlyAlaThrGlnGlyTrpAspAlaPro 300  
 DB 918 GCTACTATGATATGCACTTCTTGAAGAGTGGAGCCCACTGAGCTCTGGATGCTCTCC 977  
 QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLeuSer 320  
 DB 978 CAAGAAGTCCCTATGCTTATAGGCAACGAGTGGCTTGGCTATGACAATATCAAGAGC 1037

QY 321 PheAspIleLeuAlaGlnTrpLeuIleHisAsnLysPheGlyGlyValaMetValTrpAla 340  
 DB 1038 TTCAGTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCC 1097  
 QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360  
 DB 1098 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGGAATTTCTCTGACTTCT 1157  
 QY 361 ThrLeuLysLeuAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380  
 DB 1158 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCCTGACGTGCTTCC 1217  
 QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400  
 DB 1218 GAGCCAGTGACTACTCTCTCCA-----GGAAGTGGAGTGGGGTGGAGCTCCGGA 1268  
 QY 401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420  
 DB 1269 GGAAGCTCTGGAGGAGTGGATTTCTGTCGCAACAAAGCAGATGGCTCTTACCTGTGGCA 1328  
 QY 421 AsnAsnArgAsnAlaPheThrHisCysValaAsnGlyValThrTyrGlnGlnAsnCysGln 440  
 DB 1329 GATGACAGAAATGCTTTTGGCAGTGCATCAATGAATCACATACAGCAGCATTTGTCAA 1388  
 QY 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454  
 DB 1389 GCAGGCTTGTTTTGTATACCAAGCTGTAAATGCTGCAACTGG 1430  
 RESULT 4  
 AK007573 1530 bp mRNA linear HTC 19-JAN-2002  
 LOCUS AK007573.1 GI:12841201  
 DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810021D23:chitinase, acidic, full insert sequence.  
 ACCESSION AK007573  
 VERSION AK007573.1  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain:CS7BL/6J) 10 day old male pancreas cDNA to mRNA, clone lib-RIKEN full-length enriched mouse cDNA library clone:1810021D23.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 Carninci, P. and Hayashizaki, Y.  
 AUTHORS High-efficiency full-length cDNA cloning  
 TITLE Meth. Enzymol. 303, 19-44 (1999)  
 JOURNAL 99279253  
 MEDLINE 10349636  
 PUBMED 10349636  
 REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 AUTHORS Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 TITLE Genome Res. 10 (10), 1617-1630 (2000)  
 JOURNAL 20499374  
 MEDLINE 11042159  
 PUBMED 11042159  
 REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 AUTHORS RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 TITLE Genome Res. 10 (11), 1757-1771 (2000)  
 JOURNAL 20530913  
 MEDLINE 11076861  
 PUBMED 11076861  
 REFERENCE 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,



[illegible]

RESULT	5
B1760250	
LOCUS	957 bp mRNA linear EST 25-SEP-2001 G0304519OF1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185486 5'
DEFINITION	mRNA sequence.
ACCESSION	B1760250
VERSION	B1760250.1 GI:15751828
KEYWORDS	EST.
SOURCE	human. Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 957)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-i@mail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAMI1463 row: c column: 23 High quality sequence stop: 780.

FEATURES	SOURCE
1. <b>High Accuracy:</b> The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.	1. <b>High Accuracy:</b> The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.
2. <b>Scalability:</b> The model is designed to handle large-scale data and complex tasks, making it suitable for enterprise-level applications.	2. <b>Scalability:</b> The model is designed to handle large-scale data and complex tasks, making it suitable for enterprise-level applications.
3. <b>Interpretability:</b> The model's decisions are transparent and explainable, allowing users to understand the underlying logic and reasoning.	3. <b>Interpretability:</b> The model's decisions are transparent and explainable, allowing users to understand the underlying logic and reasoning.
4. <b>Robustness:</b> The model is highly resistant to adversarial attacks and data poisoning, ensuring reliable performance in real-world scenarios.	4. <b>Robustness:</b> The model is highly resistant to adversarial attacks and data poisoning, ensuring reliable performance in real-world scenarios.
5. <b>Efficiency:</b> The model is optimized for fast inference and low resource consumption, making it ideal for deployment on edge devices.	5. <b>Efficiency:</b> The model is optimized for fast inference and low resource consumption, making it ideal for deployment on edge devices.

stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH-MGC Library."

BASE COUNT 222 a 277 c 227 g 231 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,05e-142 Length: 957  
Score: 1438.50 Matches: 284  
Percent Similarity: 94.16% Conservative: 6  
Best Local Similarity: 92.21% Mismatches: 11  
Query Match: 57.40% Indels: 8  
DB: 13 Gaps: 1

US-10-004-219B-14 (1-455) x BI760250 (1-957)

Qy	65	LysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTirPasnPheGlyThrAlaPro	84
Db	45	CGGACAGCCAGCTGAACACTCTCTGCCATTGGAGCTGGAATCTCGGAGCTGCCCT	104
Qy	85	PheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLys	104
Db	105	TTCACTGCCATGGTTTCTACTCTGAGAACCGCAGACTTTCATCACTCAGTCATCAAA	164
Qy	105	PheLeuAArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg	124
Db	165	TTCTCGCCAGCATATGAGTTTGACGGCGTGAGCTTTCTACTGGGAGTAGTACCTGGCTCTCGT	224
Qy	125	GlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAla	144
Db	225	GGGAGCCCTCTCAGACACAGCACTCTTCACTGTCTGTGTGAGGAAATGCGTGAAGCT	284
Qy	145	PheGluGlnGluAlaLysGlnLeuAsnLysProArgLeuMetValThrAlaAlaValAla	164
Db	285	TTTGACGAGAGGCCAAGCAGATCAACAGCCAGGCTGATGGTCACTGTGCAGTAGCT	344
Qy	165	AlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnIleuSerGlnTyrLeuAsp	184
Db	345	GCTGGCATCTCCAATATCCAGTCGGCTATGAGATCCCCCACTGTCCAGTAGTACCTGGAC	404
Qy	185	TyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsn	204
Db	405	TATATCCATGTCATGACCTAGCCTCATGGCTCTCTGGGAGGCTACACTGGAGAGAAC	464
Qy	205	SerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyr	224
Db	465	AGCCCCCTCTACAAATATCCCGACTGACACCGGCAGCAACGCCCTACCTCAATGTGGATTAT	524
Qy	225	ValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPhePro	244
Db	525	GTCAATGAATCTCTGGAAGGACAATGGAGCACCACTCAGAAAGCTCATCTGTGGATTCCCT	584
Qy	245	ThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThr	264
Db	585	ACTATGGACACACTTCACTCTGAGCAACCCCTCCAACTGGAAATGTGTGCCCCACC	644
Qy	265	SerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleThrAlaTyrTyrGlu	284
Db	645	TCGTGTGCTGGTCTGTCTGGGCCCTATGCCAAGGAGTCTGGATCTGGGCTTACTACGAG	704
Qy	285	IleCysThrPheLeuLys-AsnGlyAlaThrGlnGlyTrpAspAlaProGln-GluValP	304
Db	705	ATCTGTACCTCTCTGAACAAATGGAGCCACTCAGGAGTAGGATGCCCCCTCAAGAGAGTGC	764
Qy	304	xxTyrAlaTyr-GlnGlyAsnValTrpValGlyTyr--AspAsnIleLysSerPheAsp	322
Db	765	CTTATGCCCTATTACGGGCACAGTGTGGTGTGGCTATTGACCAACATCAAGAGCTTTCGAT	824
Qy	323	IleLysAlaGlnTrpLeu-LysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAs	342

```

|||||
825 ATTAAGGCTCAATGGCTTTAGGCACACCAATTTGGAGCGCATGGTCTGGCCATTGA 884
|||||
342 Pleu-AspAspPhe-ThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
|||||
885 TCTCGATAACTTCCACTGGGCACTTCTGCAACCA-GGCAGTTTTTCCCTAATTTCCACC 943
|||||
362 LeuLysLysAla 365
|||||
944 TGTAAACCGCT 955

RESULT 6
BM923480 1105 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6623879 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5759117
DEFINITION 5', mRNA sequence.
ACCESSION BM923480
VERSION BM923480.1 GI:19373859
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1105)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsabbs@mail.nih.gov.
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M12804 row: a column: 06
High quality sequence stop: 668.
FEATURES
Location/Qualifiers
1..1105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5759117"
/lab_host="NIH MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC Library."
BASE COUNT 270 a 318 c 285 g 231 t
ORIGIN
Alignment Scores:
Pred. No.: 1,496-135 Length: 1105
Score: 1375.00 Matches: 280
Percent Similarity: 77.87% Conservative: 5
Best Local Similarity: 76.50% Mismatches: 18
Query Match: 54.87% Indels: 65
DB: 14 Gaps: 5
US-10-004-219B-14 (1-455) x BM923480 (1-1105)

65 LysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyLysThrAsnPheClyThrAlaPro 84
|||||
62 CGGACAGCCAGCTGAAACTCTCTGGGCCATTTGGAGGCTGGAACTTCGGGACTGC-CCC 120

```

```

QY 85 PheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLys 104
DB 120 -----
QY 105 PheLeuArgGlnTyrGluPheAspClyLeuAspPheAspTyrGluTyrProGlySerArg 124
DB 120 -----
QY 125 GlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAla 144
DB 121 -----GAATGCGTGAAGCT 135
QY 145 PheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAla 164
DB 136 TTTGAGCAGGAGGGCCAGCAGATCAACAGCCAGGCTGATGCTCACTGCTGCAGTAGCT 195
QY 165 AlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrIleuAsp 184
DB 196 GCTGGCATCTCAATATATCCAGTCTGGCTATGAGATCCCCCACTGTACAGTACTCTGGAC 255
QY 185 TyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsn 204
DB 256 TACATCCATGTATGACCTACGACCTCCATGGCTCTGGAGGGGTACACTGGAGAGAAC 315
QY 205 SerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyr 224
DB 316 AGCCCCCTCTACAAATATCCGACTGACACCGGACGACCTACCTCAATGTGGATTAT 375
QY 225 ValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPhePro 244
DB 376 GTCATGAATCTACTGGAGGCAATGGAGCACGAGCTGAGAAGCTCATGCTGGATTCCCT 435
QY 245 ThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThr 264
DB 436 ACCTATGGACACCACTTCATCTGAGCAACCCCTCCACACTGGNATTGGTGCCCCACC 495
QY 265 SerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGlu 284
DB 496 TCTGGTGTGCTGCTGCTGGGCCCTATGCCAAGGAGTCTGGGATCTGGGCTTACTACGAG 555
QY 285 IleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValPro 304
DB 556 ATCTGTACTCTTCTGAAATAATGGAGCCACTCAGGATGGGATGCCCTCAGGAAGTGCCT 615
QY 305 TyrAlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleLysSerPheAspIleLys 324
DB 616 TATGCTTATCAGGGCAATGTGTGGTTGGCTATGCAACATCAAGAGCTTCGATATTAG 675
QY 325 AlaGlnTyrLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAlaIleAspLeuAsp 344
DB 676 GCTCAATGGCTTAAGCACACAAATTTGGAGGGCCATGGTCTGGGCCATTGATCTGGAT 735
QY 345 AspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr-Leu-LysL 364
DB 736 GACTTCACTGGCAGCTTCTGTCAC-CAGGCGCAAGGTTCCTCCCTAATCTCCACCCCTGAAAGA 794
QY 364 ysAlaLeuGlyLeuGlnSerAlaSerCysThr-AlaProAlaGlnProIleGluProIle 383
DB 795 AGGCCCTCGGCTGCAGAGTGCAGTTCACCGGCTCCAGCTCAGCCATTGAGCCAATA 854
QY 384 ThrAlaAlaProSerGlySerGlyAsn---GlySerGlySerSerSer-----SerGly 400
DB 855 ACTGCTGCTCCAGTGGGAACCGGGAACCGGAGCGCGGAGTAGTAACAGCTCTCTGGAGGCG 914
QY 401 GlySerSerGlyGlySerGlyPheCysAlaVal---ArgAlaAsnGly---LeuTyrPro 418
DB 915 ACCTCGGGAAGCAATTGGAATTTCTGCTCTGTCAAAAGNCCACACGGGCCCTTCTTACCCC 974
QY 419 ValAlaAsnAsn 422
DB 975 GCTGGGGAANAAT 986

```

RESULT 7



```

BQ231098      892 bp      mRNA      linear      EST 02-MAY-2002
LOCUS      AGENCOURT 7578138 NCI CGAP St1 Mus musculus cDNA clone
DEFINITION      IMAGE:6051580 5', mRNA sequence.
ACCESSION      BQ231098
VERSION      BQ231098.1 GI:20412498
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      1 (bases 1 to 892)
JOURNAL      NIH-MGC http://imgc.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs@mail.nih.gov
              Tissue Procurement: Jeffrey E. Green, M.D.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM13305 row: k column: 05
              High quality sequence stop: 772.
              Location/Qualifiers
                1..892
                  /organism="Mus musculus"
                  /db_xref="taxon:10090"
                  /clone_lib="IMAGE:6051580"
                  /lab_host="NCI CGAP St1"
                  /note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
                  Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
                  Average insert size 1.77 Kb. Library constructed by Life
                  Technologies. Note: this is a NCI CGAP Library."
              BASE COUNT      219 a 226 c 216 g 230 t
              ORIGIN
                1..892
                  /organism="Mus musculus"
                  /db_xref="taxon:10090"
                  /clone_lib="IMAGE:6051580"
                  /lab_host="NCI CGAP St1"
                  /note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
                  Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
                  Average insert size 1.77 Kb. Library constructed by Life
                  Technologies. Note: this is a NCI CGAP Library."
              Alignment Scores:
                Pred. No.:      1,22e-130      Length:      892
                Score:      1327.50      Matches:      242
                Percent Similarity:      87.21%      Conservative:      17
                Best Local Similarity:      81.48%      Mismatches:      19
                Query Match:      52.97%      Indels:      19
                DB:      14      Gaps:      1
              US-10-004-219b-14 (1-455) x BQ231098 (1-892)
              QY      1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
              DB      36 TACAATCTGATGCTATTTTCCACCACTGGGCCAGTATCGCCAGGCTTGGGAGCTTC 95
              QY      21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
              DB      96 AAGCCTGATGACATTAACCCCTGCTGTACTACCTGATCTATGCTTGTGGGATG 155
              QY      41 GlnAsnAsnGluIleThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
              DB      156 CAGAACAAATGAGATCACCACTAGAAATGGAATGATGTTACTCTCTATAAAGCTTCAAT 215
              QY      61 GlyLeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyTyrPheAsn 80
              DB      216 GACTTGAAAAACAGAACGAACTGAAACCCCTCTGCAATGGAGCTGGAACTTT 275
              QY      81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
              DB      276 GGAACCTGCTCTTCACTACCATGGTTTCCACTCTCTCAGAACCGCCAGACCTTCATTACC 335
              QY      101 SerValIleIlePheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
              DB      336 TCAGTCATCAAAATTTCTGCGTCAGTATGGTTTGGAGTGGAGCTGGAGCTGGGAATAC 395

```

```

QY      121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
DB      396 CCAGGCTCAGCTGGGAGCCCTCCTCAGGACAAGCATCTCTTCACTGTCTCGTGGAGGAA 455
QY      141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
DB      456 ATGCGTGAAGCTTTTGGAGCAGGAGCTATTGAGCAACACGCCCAGACTGATGTTACT 515
QY      161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
DB      516 GCTGCTGATGCTGGTGGGATTTCCAACTCCAGGCTGCTATGAGATCCCTGAATTTCT 575
QY      181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
DB      576 AAGTACCTGGATTTTCATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 635
QY      201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220
DB      636 ACTGGGAGAAATAGTCTCTTTACAAATACCTACTAGACTGGTAGCAATGCCTACCTC 695
QY      221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
DB      696 AATGTGGATTTATGTCATGAACTATTGGAGAACATGAGCCCGCCAGCTGAGAGCTCAT 755
QY      241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
DB      756 GTTGGATTTCCAGAGTATGGACACACCTTCCTCTGAGAACCCCTCTGATTAATGGAATT 815
QY      261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
DB      816 GGTGCCCCCTACCTCTGTGTGGTGGCCCTGCTGGCCCTATACACAGA----- 860
QY      281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrp 297
DB      861 -----CAGGCTGG 869
RESULT 8
BG867815      838 bp      mRNA      linear      EST 29-MAY-2001
LOCUS      602786336F1 NCI CGAP_SG2 Mus musculus cDNA clone IMAGE:4912544 5',
DEFINITION      mRNA sequence.
ACCESSION      BG867815
VERSION      BG867815.1 GI:14218355
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      NIH-MGC http://imgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs@mail.nih.gov
              Tissue Procurement: Jeffrey E. Green, M.D.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM10816 row: g column: 09
              High quality sequence stop: 824.
              Location/Qualifiers
                1..838
                  /organism="Mus musculus"
                  /strain="FVB/N"
                  /db_xref="taxon:10090"
                  /clone="IMAGE:4912544"
                  /clone_lib="NCI CGAP_SG2"
                  /lab_host="DH10B (T1 phage-resistant)"
                  /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:

```



NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 209 a 211 c 198 g 220 t

ORIGIN

Alignment Scores:

Pred. No.: 1,47e-124 Length: 838  
Score: 1270.00 Matches: 229  
Percent Similarity: 93.18% Conservativeness: 17  
Best Local Similarity: 86.74% Mismatches: 18  
Query Match: 50.68% Indels: 0  
DB: 12 Gaps: 0

US-10-004-219B-14 (1-455) x BG867815 (1-838)

```
Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Dy 34 TACAATCTGATATGCTATTTACCAACTGGGCCAGATATCGGCCAGGTCTGGGGAGCTTC 93
Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Dy 94 AAGCTGTGATGATTAACCTCGCTGTGATCTACCTGATGATGCTTTGCTGGGATG 153
Qy 41 GlnAsnGluIleThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
Dy 154 CAGAACATGATGATCACCACCATAGATGAATGAATGATGTTACTCTCTATAAAGCTTCAAT 213
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Dy 214 GACTTGAATAAAGCAGGACAGCAAACTGAAACCTCTCTGGCAATTTGGAGGCTTGAACCTT 273
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Dy 274 GGAATGCTCTCTTTCACATACATGATGTTCCACTTCTCAGAACCCGACACCTTCAATACC 333
Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Dy 334 TCAGTCATCAATTTCTGGCTCAGTATGGTTTGATGACTGGACCTGGACTGGGAATAC 393
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Dy 394 CCAGGCTCAGCTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTCTGGTGAAGGAA 453
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Dy 454 ATGCGTGAAGCTTTTGGAGGAGGCTTATGAGACCAACAGGCCAGACTGATGTTTACT 513
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Dy 514 GCTGCTGTAGCTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACCTTCT 573
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Dy 574 AAGTACCTGGATTTTCATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 633
Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Dy 634 ACTGGGGAGATAGTGCCTTTTACAAATACCTCTACTGAGACTGGTAGCAATGCCTTACCTC 693
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
Dy 694 AATGTGGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Dy 754 GTTGATTTCCAGAGTTTGGACACACCTTCACTCTGAGAAACCCCTCTGATAATGATTT 813
Qy 261 GlyAlaProThr 264
Dy 814 GGTGCCCTTACT 825
```

RESULT 9

BI517797  
LOCUS

DEFINITION BI517797 1038 bp mRNA linear EST 29-AUG-2001  
603042049f1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5182357 5',  
mRNA sequence.

ACCESSION

VERSION BI517797.1 GI:15342589

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1038)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cagbbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L14M1455 row: a column: 14

High quality sequence start: 3

High quality sequence stop: 788.

FEATURES

source

1..1038

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5182357"

/clone\_lib="NIH\_MGC\_116"

/lab\_host="DH10B"

/note="Organ: pooled colon, kidney, stomach; Vector:

pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA

source anonymous pool of 3 colons, age 26 yo male, 49 yo

female, 71 yo male colon; 46 yo male kidney, and pool of 2

stomachs, 62 yo male and 70 yo female. Library is

oligo-dT primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.4 kb,

insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed by C.

Gruber (Invitrogen). Research Genetics tracking code

023. Note: this is a NIH MGC Library."

BASE COUNT 273 a 275 c 293 g 197 t

ORIGIN

Alignment Scores:

Pred. No.: 2,17e-122 Length: 1038

Score: 1251.00 Matches: 237

Percent Similarity: 92.69% Conservativeness: 4

Best Local Similarity: 91.15% Mismatches: 13

Query Match: 49.92% Indels: 6

DB: 13 Gaps: 2

US-10-004-219B-14 (1-455) x BI517797 (1-1038)

Qy 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221

Dy 2 GGAGAGAACAGCCCTCTACAAATACCCGACTGACACCGGCAGCACGCTACCTCAAT 61

Qy 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241

Dy 62 GTGATTATGTCATGAACACTACTGGAAGACAACTTCTATCTGAGCAACCTCCACACTGGAAATCGTT 121

Qy 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261

Dy 122 GGATTCCTACCTATGACACAACTTCTATCTGAGCAACCTCCACACTGGAAATCGTT 181

Qy 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281

Dy 182 GCCCCACCTCTGGTGTGCTGCTGCTGGGCCCTATGCCAAGGAGTCTGGGATCTGGGCT 241

282 TyrTyrGluileCysThrPheLeuLysAsnGlyValAlaThrGlnGlyTrpAspAlaProGln 301  
Db TACTACGAGATCTGTACCTCTCTGAAAAATGAGCCACTCAGGATGGGATGCCCCCTCAG 301

302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321  
Db GAAGTCCCTTATGCTTATCAGGCAATGTGTGGTGGCTATGACAACTCAGAGCTTC 361

302 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyValAlaMetValTrpAlaIle 341  
Db GATATTAGGCTCAATGGCTTAGCAACAATCTGGAGCGCCCAATGCTGGGCCAT 421

342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361  
Db GATCTGGATGACTTCACTGGCACTTCTGCAACAGCGCAAGTTTCCCTTAATCTCCACC 481

362 LeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGlu 381  
Db CTGAGAAAGCCCTCGGGCTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAG 541

382 ProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGlyGly 401  
Db CCATTAACCTGCTCTCCAGTGCAGCGGAAACGGAGCGGAGTAGCAGCTCTGGAGGC 601

402 SerSerGlyGlySerGlyPhe---CysAlaValAlaArgAlaAsnGlyLeuTyrProValAla 420  
Db AGCTCGGAGGAGGAGTGTCTTGTCTTGGCAGAGCAACAGAGCTCTAACCCCGTGGGC 661

421 Asn---AsnArgAsnAlaPheTrp-HisCysValAlaGlyValThrGln-GlnAsn- 438  
Db AAATTTACCAAGAGAGTCTCTGGGACCTGCGTGAATGGAGTGCAGTACAGCGCAGAACT 721

439 CysGlnAlaGlyLeuVal-PheAspThrSerCysAspCysCysAsnTrpAla 455  
Db TGGCAGCGCGGCTGTCTCTGAGACCAAGCTGTGAATGCTGCAACTGGGCA 773.

RESULT 10  
BQ231064  
LOCUS BQ231064 872 bp mRNA linear EST 02-MAY-2002  
DEFINITION AGENCOURT 7578230 NCI CGAP St1 Mus musculus cDNA clone  
IMAGE:6051490 5', mRNA sequence.  
BQ231064  
VERSION BQ231064.1 GI:20412464  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 872)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM13305 row: g column: 11  
High quality sequence stop: 724.  
Location/Qualifiers  
1..872  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6051490"  
/clone\_lib="NCI CGAP St1"  
/lab\_host="DH10B (Tl-resistant)"  
/note="Organ: stomach; Vector: pCMV-SPORT6; Site:1: SalI;  
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dr.

FEATURES  
source

Average insert size 1.77 kb. Library constructed by Life Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 218 a 222 c 205 g 225 t 2 others

ORIGIN

Alignment Scores:  
Pred. No.: 1.91e-121 Length: 872  
Score: 1241.00 Matches: 229  
Percent Similarity: 92.45% Conservative: 16  
Best Local Similarity: 86.42% Mismatches: 18  
Query Match: 49.52% Indels: 2  
DB: 14 Gaps: 0

US-10-004-219B-14 (1-455) x BQ231064 (1-872)

QY 1 TyrGlnLeuTyrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
Db 68 TACAATCTGATATGCTATTTACCAACTGGGCCAGTATCGGCCAGGCTCTGGGAGCTTC 127

QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40  
Db 128 AAGCCTGATGACATTAAACCCCTGCTGTGTACTACCTGATCTATGCTTCTGGGATG 187

QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60  
Db 188 CAGAAACAATGATGATCACCACCATAGATGAATGATGTTACTCTCTATAAAGCTTTCAAT 247

QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80  
Db 248 GACTTGAATAACAGGAACAGCAACTGAACCCCTCTGGCAATTTGGAGGCTGGAACTTT 307

QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
Db 308 GGAACCTGCTCTTTTCACTACCATGCTTTTCCACTTCTCAGAACCCGACAGCTTCAATACC 367

QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120  
Db 368 TCAGTCATCAAAATTTCTCGCTCAGTATGGGTTGTGATGACTGGACCTGGACTGGGAATAC 427

QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
Db 428 CCAGGCTCAGCTGGGAGCCCTCTCCAGGACAAAGCATCTCTTCACTGTCTGGTGAAGNA 487

QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
Db 488 ATGCTGAAGCTTTTGGAGCAGGAGCTATTGAGAGCAACAGGCCCCAGACTGATGTTACT 547

QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
Db 548 GCTGCTGTAGCTGGTGGGATTTCCAAACATCCAGGCTGCTATGAGATCCCTGAACCTTCT 607

QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200  
Db 608 AAGTACCTGGATTTTCATCAATGTCATGACATATGACCTCCATGGCTCTGGAGGCTAC 667

QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
Db 668 ACTGGGGAGATAGTCTCTTTTACAAATACCCTACTGAGACTGGTAGCAATGCCCTACCTC 727

QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyValProAlaGlyLysLeuIle 240  
Db 728 AATGTGGATTATGTCATGAACCTATTGGGAAGAACAAATGGAGCCCACTGAGAAGCTCAAT 787

QY 241 ValGlyPheProThrTyrGlyHisAsn-PheIleLeuSerAsnProSerAsnThr-GlyI 260  
Db 788 GTTGGATTTCCAGAGTAGGACACACCCCTTCATCTCTGAGAAACCCCTCTGTATATGGNA 847

QY 260 LeGlyAlaPro 263  
Db 848 TTGGTGGCCCC 858

RESULT 11  
BG866417

LOCUS BG866417 809 bp mRNA linear EST 29-MAY-2001  
DEFINITION 602785352F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4911421 5',  
mRNA sequence.  
ACCESSION BG866417  
VERSION BG866417.1 GI:14216957  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 809)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LHAM0813 row: h column: 14  
High quality sequence stop: 804.  
Location/Qualifiers  
1..809  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4911421"  
/clone\_lib="NCI CGAP SG2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 193 a 200 c 208 g 207 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,536-121 Length: 809  
Score: 1237.00 Matches: 223  
Percent Similarity: 90.71% Conservative: 21  
Best Local Similarity: 82.90% Mismatches: 24  
Query Match: 49.36% Indels: 1  
DB: 12 Gaps: 0  
US-10-004-219B-14 (1-455) x BG866417 (1-809)  
QY 71 ThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrAlaProPheThrAlaMetValSer 90  
DB 2 ACCCTCTGGCAATGGAGCGTGGAACTTGAACCTGCTTCCTTACCTACCATGTTTCC 61  
QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlu 110  
DB 62 ACTTCTCAGACCGCCAGACCTTCATTACCTCAGTCATCAATTTCTGGTCAGATGGG 121  
QY 111 PheAspGlyLeuAspPheAspTrpGlnTyrProGlySerArgGlySerProProGlnAsp 130  
DB 122 TTTGATGACTGGACCTGGACTGGGAATACCCAGGCTCAGCTGGGAGCCCTCTCAGGAC 181  
QY 131 LysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAlaLys 150  
DB 182 AMGATCTCTTCTCTCTGCTGGTGAAGGAATGCGTGAAGCTTTTGAGCAGGAGCTATT 241  
QY 151 GlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsnIle 170  
DB 242 GAGACACAGCCAGCCAGACTGATGTTACTGCTGTAGCTGGTGGGATTTCCACATC 301  
QY 171 GlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMetThr 190  
DB 302 CAGGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGATTTCATCATGTCATGACA 361

QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrIlystYr 210  
DB 362 TATGACCTCCATGGCTCTCTGGAGGGCTACTCTGGGAGATAGTCTCTTTTACAAATAC 421  
QY 211 ProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrpLys 230  
DB 422 CCTACTGAGACTGGTAGCAATGCTTACCTCAATGCGATTATGTCATGAACCTATTGGGAG 481  
QY 231 AspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHisAsnPhe 250  
DB 482 AACATGGAGCCCGCAGCTGAGAAGCTCAATTGTGATTCCAGAGTATGGACACACCTTC 541  
QY 251 IleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyValaGlyProAla 270  
DB 542 ATCTGTGAGAAACCCCTCTGATAATGGAATGGTGGCCCTACTCTGTGTATGGCCCTGCT 601  
QY 271 GlyProTyrAlaLysGluSerGlyIleTrpAlaTyrTyrGluIleCysThrPheLeuLys 290  
DB 602 GGGCCCTATACGACAGACAGCTGGGTTCTGGGCTACTATGAGATNTGCACTTTCTGAGA 661  
QY 291 AsnGlyAlaThrGlnGlyTrp-AspAlaProGlnGluValProTyrAlaTyrGlnGlyAs 310  
DB 662 AGTGAGGCCACTGAGGTCTGGGATGCTCCCAAGAGTGGCCCTATGCTATAGGGCCAA 721  
QY 310 nValTrpValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTrpLeuLysHi 330  
DB 722 CGAGTGGCTTGGCTATGACATTATCAGAGCTTCAGTGTAAAGGTCAGTGGCTTAAGCA 781  
QY 330 sAsnLysPheGlyGlyAlaMetVal 338  
DB 782 GAACAATTTGGAGTGCCATGATC 806  
RESULT 12  
BG866304 764 bp mRNA linear EST 29-MAY-2001  
LOCUS 602788819F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4920173 5',  
mRNA sequence.  
ACCESSION BG866304  
VERSION BG866304.1 GI:14219844  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 764)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LHAM0836 row: e column: 06  
High quality sequence stop: 760.  
Location/Qualifiers  
1..764  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4920173"  
/clone\_lib="NCI CGAP SG2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 198 a 195 c 180 g 191 t

## ORIGIN

Alignment Scores:  
Pred. No.: 3,97e-118 Length: 764  
Score: 1209.00 Matches: 219  
Percent Similarity: 94.35% Conservative: 15  
Best Local Similarity: 88.31% Mismatches: 14  
Query Match: 48.24% Indels: 0  
DB: 12 Gaps: 0

US-10-004-219B-14 (1-455) x BG869304 (1-764)

QY 8 ThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMetProAspAsnIleAspPro 27  
DB 13 ACCAATGGGCGCCAGTATCGCCAGGCTTGGGAGCTTCAAGCTGATGACATTAACCC 72  
QY 28 CysLeuYerThrHisLeuIleTyrAlaPheAlaGlyArgGlnAsnGlnIleThr 47  
DB 73 TGCCTGTGTACTCACCTGATCTATGCTTTTGTGGGATGCAGAACATGAGATCACCC 132  
QY 48 IleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGlyLeuLysAsnSer 67  
DB 133 ATAGATGAATGAATGTTACTCTCTATAGCTTTCAATGACTTGAATAACAGACAC 192  
QY 68 GlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPheThrAla 87  
DB 193 AAACGAAACCCCTCTGGCAATTTGAGGCTGGAACCTTTGGAACCTCTCTTCACTACC 252  
QY 88 MetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArg 107  
DB 253 ATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATTAACCTCAGTCATCAATTTCTGCT 312  
QY 108 GlnTyrGluPheAspGlyLeuAspPheAspTyrGluTyrProGlySerArgGlySerPro 127  
DB 313 CAGTATGGGTTTGTATGGACTGGACCTGGAGTGGGAATACCCAGGCTCACTGGGAGCCCT 372  
QY 128 ProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGln 147  
DB 373 CCTCAGCAACAGCATCTCTTCACTGCTCTGCTGTTGAAGAAATCGTGAAGCTTTTGAGCAG 432  
QY 148 GluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaGlyIle 167  
DB 433 GAGGCTATTGAGAGCAACAGGCCAGACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 492  
QY 168 SerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHis 187  
DB 493 TCAACATCCAGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGATTTTCATCCAT 552  
QY 188 ValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyLysAsnSerProLeu 207  
DB 553 GTCATGACATATGACCTCCATGCTCTCTGGAGGCTACACTGGGAGAAATAGTCTCTT 612  
QY 208 TyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsn 227  
DB 613 TACAAATACCTACTAGAGCTGTGTAGCAATGCCATCTCAATGTGATTTATGTGATGAAC 672  
QY 228 TyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGly 247  
DB 673 TATTGGAAGAACAAATGGAGCCCGCCAGCTGAGAGCTCATTTGTTGATTCACAGATGGA 732  
QY 248 HisAsnPheIleLeuSerAsnPro 255  
DB 733 CACACCTTTCATCTGAGAAACCT 756

## RESULT 13

BG872647  
LOCUS 763 bp mRNA linear EST 29-MAY-2001  
DEFINITION 602793705F1.NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4924897 5',  
mRNA sequence.  
ACCESSION BG872647  
VERSION BG872647.1 GI:14223187  
KEYWORDS EST.  
SOURCE house mouse.

## ORGANISM

## Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 763)

NIH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLNMI0848 row: j column: 02

High quality sequence stop: 760.

## FEATURES

## source

1..763

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4924897"

/lab host="NCI CGAP\_SG2"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:

NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dT. Average insert size 1.3 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 199 a 190 c 178 g 196 t

## ORIGIN

Alignment Scores:  
Pred. No.: 6,47e-118 Length: 763  
Score: 1207.00 Matches: 222  
Percent Similarity: 93.68% Conservative: 15  
Best Local Similarity: 87.75% Mismatches: 16  
Query Match: 48.16% Indels: 1  
DB: 12 Gaps: 0

US-10-004-219B-14 (1-455) x BG872647 (1-763)

QY 3 LeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMetPro 22  
DB 1 CTGATATGCTATTTTCAACCACTGGGCCAGTATGCCAGGCTCGGGAGCTTCAAGCT 60  
QY 23 AspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGlnAsn 42  
DB 61 GATGACATTAACCCCTGCTGTACTCACCTGATCTATGCTTGTGGATGAGAAC 120  
QY 43 AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGlyLeu 62  
DB 121 AATGAGATCACCAACATAGAAATGGAATGATGTTACTCTCTATAAGCTTTCAATGACTTG 180  
QY 63 LysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr 82  
DB 181 AAAACAGACAGACAGCAAACTGAAACCTCTCTGGCAATGGAGGCTGGAACCTTGGAACT 240  
QY 83 AlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerVal 102  
DB 241 GCTCCTTTTCACTTACCATGGTTTCCACTTCTCAGAACCCGACACCTTCAATCACTCAGTC 300  
QY 103 IleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrProGly 122  
DB 301 ATCAAAATTTCTGCGTCAGTATGGGTTTGTATGACTGGACCTGGAGTGGAAATACCAAGC 360  
QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArg 142  
DB 361 TCACGTGGGAGCCCTCTCTCAGACAAAGCATCTCTTCACTGCTCTGCTGGTGAAGAAATGCT 420  
QY 143 GluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAla 162

421 GAAGCTTTTGGAGCAGGCTATTGAGAGCAACAGGCCCCAGACTGATGGTTACTGCTGCT 480  
 163 ValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyr 182  
 481 GTAGCTGGTGGATTCACCAATCCAGCTGGCTATGAGATCCCTGAAACTTCTAAGTAC 540  
 183 LeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGly 202  
 541 CTGGATTTCATCCATGTCATGACATATGACCTCCATGCTCTCTGGAGGGCTACACTGG 600  
 203 GluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnVal 222  
 601 GAGAACTGCTCTTTCACAAATACCTTACTGAGACTGGTAGCAATGGCTTACCTCAATGT 660  
 223 AspTyrValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIleValGly 242  
 661 GATTATGTCATGAACTATTGGNAGAACATGGAGC-CCAGCTGAGNAGCTCATTTGTGA 719  
 243 PheProThrTyrGlyHisAsnPheIleLeuSerAsnPro 255  
 720 TTCCAGAGTATGGACACACCTTTCATCTCTGAGAAACCT 758

RESULT 14  
 BG868741  
 LOCUS  
 DEFINITION 602787570F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4913800 5', mRNA sequence.  
 ACCESSION BG868741  
 VERSION BG868741.1 GI:14219281  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 TITLE NIH-MGC http://mgi.nci.nih.gov/.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov  
 Plate: LHAM10819 row: k column: 17  
 High quality sequence stop: 836.

FEATURES  
 source  
 1..837  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4913800"  
 /clone\_lib="NCI CGAP SG2"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 212 a 211 c 197 g 217 t

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.22e-117 Length: 837  
 Score: 1205.00 Matches: 219  
 Percent Similarity: 92.58% Conservative: 18  
 Best Local Similarity: 85.55% Mismatches: 18  
 Query Match: 48.08% Indels: 1  
 DB: 12 Gaps: 0

US-10-004-219B-14 (1-455) x BG868741 (1-837)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe 20  
 Db 66 TACAATCTGATATGCTATTTTACCACCACTGGGCCAGTATCGCCAGGTCCTGGGAGCTTC 125  
 QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40  
 Db 126 AAGCCTGATGACATTAACCCCTGCTGTGTACTCACCTGATCTATGCTTGTGGGATG 185  
 QY 41 GlnAsnGluIleThrThrIleGluTyrAsnAspValThrLeuTyrGlnAlaPheAsn 60  
 Db 186 CAGAACATGAGATCACCACTAGATAGATGGATGATTTACTTCTATPAAGCTTTTCAAT 245  
 QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80  
 Db 246 GACTTGAACACAGAACAGCAACTGAAACCCCTCTCGCAATTGGAGGCTGGAACCTT 305  
 QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100  
 Db 306 GGAATCTGCTCTTTCACCTACCATGCTTTCCACTTCTCAGAACCCGACAGCTTCATTACC 365  
 QY 101 SerValIleLysPheLeuAtgGlnTyrCluPheAspGlyLeuAspPheAspTyrGluTyr 120  
 Db 366 TCAGTCATCAAAATTTCTGCGTCAGTATGGGTTTGATGACTGGACCTGGAGCTGGGAATAC 425  
 QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140  
 Db 426 CCAGGCTCAGCTGGGAGCCCTCTCCAGGACCAAGCATCTTTCACCTGCTCTGGTGAAGGA 485  
 QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160  
 Db 486 ATGCGTGAAGCTTTTGGAGCAGGCTATTGAGAGCAACAGGCCCCAGACTGATGTTACT 545  
 QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180  
 Db 546 GCTGCTGATGCTGGTGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCT 605  
 QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyr 200  
 Db 606 AAGTACCTGGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTGGGAGGGCTAC 665  
 QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220  
 Db 666 ACTGGGAGATAGTCTCTTTACAAATACCTACTGAGACTGGTAGCAATGCCCTACCTC 725  
 QY 221 AsnValAspTyrValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIle 240  
 Db 726 AATGTGGATTATGTCATGAACATTTGGGAGAACAAATTTGGAGCCCGCCAGTGAAGCTCAT 785  
 QY 241 ValGlyPheProThrTyrGlyHisAsnPhe-IleLeuSerAsnPro 255  
 Db 786 GTTGGATTTCCAGAGTATGGACACACCTTTCATCTCTGAGAAACCT 831

RESULT 15  
 BQ219632  
 LOCUS  
 DEFINITION AGNCOURT 7572993 NCI CGAP\_St1 Mus musculus cDNA clone  
 IMAGE:6051913 5', mRNA sequence.  
 ACCESSION BQ219632  
 VERSION BQ219632.1 GI:20401032  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 TITLE NIH-MGC http://mgi.nci.nih.gov/.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LAM13306 row: i column: 02  
High quality sequence stop: 654.

## FEATURES

source

1..939

/organism="Mus musculus"

/db\_xrefs="taxon:10090"

/clone\_image="6051913"

/lab\_host="DH10B (T1-resistant)"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.77 kb. Library constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 232 a 246 c 224 g 237 t

ORIGIN

## Alignment Scores:

Pred. No.:	3,06e-117	Length:	939
Score:	1202.00	Matches:	233
Percent Similarity:	87.41%	Conservative:	17
Best Local Similarity:	81.47%	Mismatches:	29
Query Match:	47.96%	Indels:	7
DB:	14	Gaps:	4

US-10-004-219B-14 (1-455) x BQ219632 (1-939)

```
QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 79 TACATCTGATATGCTATTTTCCACCACTGGGCCAGTATCGGCCAGGTCTGGGAGCTTC 138
QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
DB 139 AAGCCTGATGACATTAACCCCTCGCTGTCTACTCACCTGATCTATGCTTGTCTGGATG 198
QY 41 GlnAsnGlnIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
DB 199 CAGAACTGATGACATCACCACTAGATGAATGATGTTACTCTCTATAAAGCTTTCAAT 258
QY 61 GlyLeuLeuAsnIleAsnSerGlnLeuLeuThrLeuAlaIleGlyTyrAsnPhe 80
DB 259 GACTTGAAGAAACAGGACAGCAACTGAAACCCCTCTGGCAATTGGAGCTGGAACTTT 318
QY 81 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
DB 319 GGAACCTGCTCTTCACTACCATGTTCCACTTCTCAGAACCGCCAGACCTTCATTACC 378
QY 101 SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
DB 379 TCAGTCATCAAAATTTTCGCTCAGTATGGTTTGTAGGACTGGACCTGGGAAATAC 438
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
DB 439 CCAGGCTCACGTGGGAGCCCTCTCAGGACAGCATCTCTTCACTGCTCTGTTGAGGAA 498
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
DB 499 ATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGTTACT 558
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
DB 559 GCTGCTGTAGCTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACCTTCC 618
QY 181 -GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTy 200
DB 619 TAAGTACCTGGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTGGGAGGCTA 678
QY 200 rThrGlyGluAsnSerProLeuTyrIleThrProThrAspThrGlySerAsnAlaTyrIle 220
```

```
DB 679 CCTGGGGAGATAGTCTCTTTTACAAATACCTACTGAGACTGGTAGCAATGCCTACTCT 738
QY 220 uAsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAla-ProAlaGluLysLeuI 240
DB 739 CAATGTGGGATATGTCATGAACACTATTGGAAAAACATGTGAGCCCCCAGCTGAGAAGCTCA 798
QY 240 leValGlyPheProThrTyrGly---HisAsnPheIleLeuSerAsnProSerAsnThrG 259
DB 799 TTGTTGGATTCCCCAAATATGGAACACACACCTTCTATCTCGAAAAAACCCCTTCTAATAATG 858
QY 259 ly-IleGlyAlaPro---ThrSerGlyAlaGlyPro---AlaGlyPro---TyrAlaLys 275
DB 859 GAAATTTGGGGGCCCCCTTACCTCTGGAGATGGCCCCCTGTGGGGCCCCCAATACCCAAAA 918
QY 276 GluSerGlyIle 279
DB 919 AGGCGTTGGGTT 930

Search completed: June 30, 2003, 01:08:01.
Job time : 1512.62 secs
```